On the complexity of the model checking problem

Florent R. Madelaine [*]
Université d'Auvergne
florent.madelaine@udamail.fr

Barnaby D. Martin [†]
Laboratoire d'Informatique de l'École Polytechnique
barnabymartin@gmail.com

Abstract

The model checking problem for various fragments of first-order logic has attracted much attention over the last two decades: in particular, for the fragment induced by $\exists$ and $\wedge$ and that induced by $\forall, \exists$ and $\wedge$, which are better known as the *constraint satisfaction problem* and the *quantified constraint satisfaction problem*, respectively. These two fragments are in fact the only ones for which there is currently no known complexity classification. All other syntactic fragments can be easily classified, either directly or using Schaefer's dichotomy theorems for SAT and QSAT, with the exception of the *positive equality free fragment* induced by $\exists, \forall, \wedge$ and $\vee$. This outstanding fragment can also be classified and enjoys a *tetrachotomy*: according to the model, the corresponding model checking problem is either tractable, NP-complete, co-NP-complete or Pspace-complete. Moreover, the complexity drop is always witnessed by a generic solving algorithm which uses quantifier relativisation (for example, in the co-NP-complete case, the model has a constant e to which all $\exists$ quantifiers may be relativised). Furthermore, its complexity is characterised by algebraic means: the presence or absence of specific *surjective hyper-operations* among those that preserve the model characterise the complexity. Our classification methodology relies on this suitably tailored algebraic approach and it suffices to classify the complexity of a finite number of cases for each model size: each case corresponds to an element of the **finite** lattice of *down-closed monoids of surjective hyper operations*. This is unlike the constraint satisfaction problem where the corresponding lattices are uncountable and essentially uncharted in general. Though finite, the number of elements of the lattice of down-closed monoids of surjective hyper operations grows rapidly as the size n of the model increases. We are able to compute suitable parts of this lattice by hand for $n = 2$ and 3, and in a computer assisted manner for $n = 4$. For arbitrarily large n, one can restrict the classification to specific monoids of surjective hyper-operations which corresponds to certain cores, for a suitable notion of core for the positive equality-free fragment of first order logic. These specific monoids enjoy a nice normal form which means that we are able to provide generic hardness proofs which mimic cases encountered when $n \leq 4$.

Keywords: Constraint Satisfaction, Galois Connection, Logic in Computer Science, Quantified Constraint Satisfaction, Universal Algebra.

1   Introduction

The *model checking problem* over a logic $\mathscr{L}$ takes as input a structure $\mathscr{D}$ and a sentence φ of $\mathscr{L}$, and asks whether $\mathscr{D} \models \varphi$. The problem can also be parameterised, either by the sentence φ, in which case the input is simply $\mathscr{D}$, or by the model $\mathscr{D}$, in which case the input is simply φ. Vardi has studied the complexity of this problem, principly for logics which subsume FO [Var82]. He describes the complexity of the unrestricted problem as the *combined complexity*, and the complexity of the parameterisation by the sentence (respectively, model) as the *data complexity* (respectively, *expression complexity*). For the majority of his logics, the expression and combined complexities are comparable, and are one exponential higher than the data complexity.

In this paper, we will be interested in taking syntactic fragments $\mathscr{L}$ of FO, induced by the presence or absence of quantifiers and connectives, and studying the complexities of the parameterisation of the

model checking problem by the model $\mathscr{D}$, that is the expression complexities for certain $\mathscr{D}$. When $\mathscr{L}$ is the *primitive positive* fragment of FO, $\{\exists, \wedge\}$-FO, the model checking problem is equivalent to the much-studied *constraint satisfaction problem* (CSP). The parameterisation of this problem by the model $\mathscr{D}$ is equivalent to what is sometimes described as the *non-uniform* constraint satisfaction problem, CSP($\mathscr{D}$) [KV00]. It has been conjectured [BKJ00,FV98] that the class of CSPs exhibits *dichotomy* – that is, CSP($\mathscr{D}$) is always either in P or is NP-complete, depending on the model $\mathscr{D}$. This is tantamount to the condition that the expression complexity for $\{\wedge, \exists\}$-FO on $\mathscr{D}$ is always either in P or is NP-complete. While in general this conjecture remains open, it has been proved for substantial classes and various methods, combinatorial (graph-theoretic), logical and universal-algebraic have been brought to bear on this classification project, with many remarkable consequences. Schaefer was a precursor and provided a dichotomy for Boolean structures using a logico-combinatorial approach [Sch78]. Further dichotomies were obtained: *e.g.* for structures of size at most three [Bul06], for undirected graphs [HN90], smooth digraphs [BKN09]. A conjectured delineation for the dichotomy was given in the algebraic language in [BJK05].

When $\mathscr{L}$ is *positive Horn*, $\{\exists, \forall, \wedge\}$-FO, the model checking problem is equivalent to the well-studied *quantified constraint satisfaction problem* (QCSP). No overarching polychotomy has been conjectured for the non-uniform QCSP($\mathscr{D}$), although the only known attainable complexities are P, NP-complete and Pspace-complete. Schaefer announced a dichotomy in the Boolean case [Sch78] between P and Pspace-complete in the presence of constants, a dichotomy which was proved to hold even when constants are not present [Dal97, CKS01]. Some partial classification were obtained, algebraically [Che04, Che08, BBC$^+$09] or combinatorially [MM06, Mar11]. A conjecture delineating the border between NP and Pspace-complete was recently ventured by Chen in the algebraic language for structures with all constants [Che12].

Owing to the natural duality between $\exists, \vee$ and $\forall, \wedge$, we consider also various dual fragments. For example, the dual of $\{\exists, \wedge\}$-FO is *positive universal disjunctive* FO, $\{\forall, \vee\}$-FO. It is straightforward to see that this class of expression complexities exhibits dichotomy between P and co-NP-complete if, and only if, the class of CSPs exhibits dichotomy between P and NP-complete. Table 1 summarises known results regarding the complexity of the model checking for syntactic fragments of first-order logic, up to this duality.

In the case of primitive positive logic, it makes little difference whether or not equality is allowed, that is the expression complexities for $\{\exists, \wedge\}$-FO and $\{\exists, \wedge, =\}$-FO are equivalent. This is because equality may be propagated out in all but trivial instances. The same is true for positive Horn logic, but is not true, e.g., for positive universal disjunctive FO. Indeed, a classification of the expression complexities over $\{\forall, \vee\}$-FO is equivalent to the unproven CSP dichotomy conjecture, though we are able to give a full dichotomy for the expression complexities over $\{\forall, \vee, =\}$-FO. The reason for this is that the equality relation in the latter simulates a disequality relation in the former. If the model $\mathscr{D}$ has $k \geq 3$ elements then $\{\exists, \wedge, \neq\}$-FO can simulate k-colourability; and, otherwise we have a Boolean model and Schaefer's dichotomy theorem provides the classification. A similar phenomenon occurs at a higher level when $\forall$ is also present.

Other fragments can be easily classified, as the model checking problem is always hard except for pathological and rather trivial models, with the notable exception of *positive equality-free first-order logic* $\{\exists, \forall, \wedge, \vee\}$-FO. For this outstanding fragment, the corresponding model checking problem can be seen as an extension of QCSP in which disjunction is returned to the mix. Note that the absence of equality is here important, as there is no general method for its being propagated out by substitution. Indeed, we will see that evaluating the related fragment $\{\exists, \forall, \wedge, \vee, =\}$-FO is Pspace-complete on any structure $\mathscr{D}$ of size at least two.

The case of $\{\exists, \forall, \wedge, \vee\}$-FO is considerably richer than all other cases (as seen on Table 1) – with the exception of $\{\exists, \wedge\}$-FO and $\{\exists, \forall, \wedge\}$-FO which are still open and active fields of research – and

Fragment	Dual	Classification?
$\{\exists, \vee\}$ $\{\exists, \vee, =\}$	$\{\forall, \wedge\}$ $\{\forall, \wedge, \neq\}$	Trivial (in L).
$\{\exists, \wedge, \vee\}$ $\{\exists, \wedge, \vee, =\}$	$\{\forall, \wedge, \vee\}$ $\{\forall, \wedge, \vee, \neq\}$	Trivial (in L) if the core of $\mathscr{D}$ has one element and NP-complete otherwise.
$\{\exists, \wedge, \vee, \neq\}$	$\{\forall, \wedge, \vee, =\}$	Trivial (in L) if $|D| = 1$ and NP-complete otherwise.
$\{\exists, \wedge\}$ $\{\exists, \wedge, =\}$	$\{\forall, \vee\}$ $\{\forall, \vee, \neq\}$	**CSP dichotomy** conjecture: P or NP-complete.
$\{\exists, \wedge, \neq\}$	$\{\forall, \vee, =\}$	Trivial if $|D| = 1$; in P if $|D| = 2$ and $\mathscr{D}$ is affine or bijunctive; and, NP-complete otherwise.
$\{\exists, \forall, \wedge\}$ $\{\exists, \forall, \wedge, =\}$	$\{\exists, \forall, \vee\}$ $\{\exists, \forall, \vee, \neq\}$	a **QCSP trichotomy** should be conjectured: P, NP-complete, or Pspace-complete.
$\{\exists, \forall, \wedge, \neq\}$	$\{\exists, \forall, \vee, =\}$	Trivial if $|D| = 1$; in P if $|D| = 2$ and $\mathscr{D}$ is affine or bijunctive; and, Pspace-complete otherwise.
$\{\forall, \exists, \wedge, \vee\}$		**Positive equality free tetrachotomy**: P, NP-complete, co-NP-complete or Pspace-complete
$\{\neg, \exists, \forall, \wedge, \vee\}$		Trivial when $\mathscr{D}$ contains only trivial relations (empty or all tuples, and Pspace-complete otherwise.
$\{\forall, \exists, \wedge, \vee, =\}$ $\{\forall, \exists, \wedge, \vee, \neq\}$ $\{\neg, \exists, \forall, \wedge, \vee, =\}$		Trivial when $|D| = 1$, Pspace-complete otherwise.

Table 1: Complexity of the model checking according to the model for syntactic fragments of FO (L stands for logarithmic space, P for polynomial time, NP for non-deterministic polynomial time, co-NP for its dual and Pspace for polynomial space).

is the main contribution of this paper. We undertook the study of the complexity of the model checking of $\{\exists, \forall, \wedge, \vee\}$-FO through the algebraic method that has been so fruitful in the study of the CSP and QCSP [Sch78,JCG97,Bul06,BBC$^+$09,Che08]. To this end, we defined *surjective hyper-endomorphisms* and used them to define a Galois connection that characterises definability under $\{\exists, \forall, \wedge, \vee\}$-FO and prove that it suffices to study the complexity of problems associated with the closed sets of the associated lattice, the so-called *down-closed monoids of unary surjective hyper-operations* (DSM for short) [MM12a]. Unlike the case of CSP where the corresponding lattice, the so-called *clone lattice*, is infinite and essentially uncharted when the domain size exceeds two, our lattice of DSMs is finite for any fixed domain. This has meant that we were able to compute the lattice for modest domain sizes, or charter parts relevant to our classification project, whether by hand for a domain of up to three elements [MM09], or using a computer for up to four elements [MM10]. These papers culminate in a full classification – a tetrachotomy – as $\mathscr{D}$ ranges over structures with up to four elements domains. Specifically, the problems $\{\exists, \forall, \wedge, \vee\}$-FO($\mathscr{D}$) are either in L, are NP-complete, are co-NP-complete or are Pspace-complete. It is a pleasing consequence of our algebraic approach that we can give a quite simple explanation to the delineation of our subclasses. A drop in complexity arises precisely when we may relativise w.l.o.g. all quantifiers of one type to a single domain element: for example, all existential quantifiers may be fixed to a special domain element e, resulting in a natural complexity drop from Pspace to co-NP. Moreover, for membership of L, NP and co-NP, it is proved in [MM12a] that it is sufficient that $\mathscr{D}$ has certain special surjective hyper-endomorphisms. For our previous example, we would have a surjective hyper-operation, i.e. a function f from D to the power set of D, such that $e \in f(d)$ for any element d of D, which is a surjective hyper-endomorphism of $\mathscr{D}$. Intuitively, a winning strategy for the existential player for some input sentence φ may be transformed through "application of f" into

3

a winning strategy where any existential variable is played on the constant e. The converse, that it is necessary to have these special surjective hyper-endomorphisms, is more subtle and was initially only an indirect consequence of our exploration of the lattice of DSMs. We settled this converse direction and the tetrachotomy for any domain size via the introduction of the novel notion of U-X-core [MM11] which is the analog for $\{\exists, \forall, \wedge, \vee\}$-FO of the core, so useful in the case of $\{\exists, \wedge\}$-FO and CSP.

The well-known notion of the core of $\mathscr{D}$ may be seen as the minimal induced substructure $\mathscr{D}' \subseteq \mathscr{D}$ such that $\mathscr{D}'$ and $\mathscr{D}$ agree on all primitive positive sentences. Equivalently, the domain D' of $\mathscr{D}'$ is minimal such that any primitive positive sentence is true on $\mathscr{D}$ iff it is true on $\mathscr{D}$ with all (existential) quantifiers relativised to D'. Cores are minimal structures in their equivalence classes, given by the equivalence relation of satisfying the same primitive positive sentences. Cores are very robust, for instance, being unique up to isomorphism, and sitting as induced substructures in all other structures in their equivalence class. A similar notion to core exists for the QCSP, but it is not nearly so robust (they need no longer be uniquely minimal in size nor sit as an induced substructure in other structures in their equivalence class [MM12b]). For the problems $\{\exists, \forall, \wedge, \vee\}$-FO$(\mathscr{D})$, a notion of core returns, and it is once again robust. The U-X-core of $\mathscr{D}$ consists of a minimal substructure induced by the union $U \cup X$ of two minimal sets U and X of D such that a positive equality-free sentence is true on $\mathscr{D}$ iff it is true on $\mathscr{D}$ with the universal quantifiers relativised to U and the existential quantifiers relativised to X. Analysing U-X-cores gives us the necessary converse alluded to in the previous paragraph. In the Pspace-complete case, some completion of the U-X-core is either fundamentally very simple and can be classified as in a two-element domain, known from [MM09], or it is a generalisation of one of the four-element cases from [MM10]. For the NP-complete and co-NP-complete cases, some completion of the U-X-core is fundamentally very simple and can be classified as an easy generalisation of a three-element domain.

We are able therefore to give the delineation of our tetrachotomy by two equivalent means. Firstly, by the presence or absence of certain special surjective hyper-endomorphisms, the so-called *A-shops* and *E-shops* (in our running example above f is an E-shop, its dual i.e. a surjective hyper-operation g such that there exists a constant u such that $g(u) = D$ would be an A-shop). Secondly, by the existence or not of trivial sets for the relativisation of universal and existential quantifiers (see Table 2). Thus, $\{\exists, \forall, \wedge, \vee\}$-FO$(\mathscr{D})$ is in L iff $\mathscr{D}$ has both an A-shop and an E-shop for surjective hyper-endomorphism, iff there exist singleton sets U and X such that a sentence of positive equality-free logic is true on $\mathscr{D}$ exactly when it is true on $\mathscr{D}$ with the universal quantifiers and existential quantifiers relativised to U and X, respectively. Otherwise, and in a similar vein, $\{\exists, \forall, \wedge, \vee\}$-FO$(\mathscr{D})$ is NP-complete (resp., co-NP-complete) if it has an A-shop (resp., E-shop) for a surjective hyper-endomorphism, iff there exists a singleton set U (resp., X) such that a sentence of positive equality-free logic is true on $\mathscr{D}$ exactly when it is true on $\mathscr{D}$ with the universal quantifiers relativised to U (resp., the existential quantifiers relativised to X). In all remaining cases, $\{\exists, \forall, \wedge, \vee\}$-FO$(\mathscr{D})$ is Pspace-complete, and $\mathscr{D}$ has neither an A-shop nor an E-shop as a surjective hyper-endomorphism, and there are no trivial sets U nor X affording the required relativisation properties.

The paper is organised as follows. In Section 2, we prove preliminary results, in Section 3, we classify the complexity of the model checking problem for all fragments (other than those corresponding to CSP and QCSP) but $\{\exists, \forall, \wedge, \vee\}$-FO. In Section 4, we classify the complexity of the model checking for $\{\exists, \forall, \wedge, \vee\}$-FO.

In more detail, in § 2.2, we present our methodology to tackle systematically the complexity of the model checking problem, discuss duality and the fragments it will suffice to classify. In § 2.3, we recall the notion of containment, equivalence and cores and extend it abstractly to any fragment $\mathscr{L}$. In § 2.5, we introduce the notion of hyper-operations and hyper-morphisms which arise naturally in the context of equality-free fragments. In § 2.6, we investigate containment, equivalence and core for $\{\exists, \forall, \wedge, \vee, \neg\}$-FO. In § 2.7, we characterise containment for $\{\exists, \forall, \wedge, \vee\}$-FO; in § 2.8, we introduce the notion of a U-X-core; highlight the link with relativisation in § 2.9; prove some basic properties of

Tetrachotomy for $\{\exists, \forall, \wedge, \vee\}$-FO($\mathscr{D}$)						
Case	Complexity	A-shop	E-shop	U-X-core	Relativises into	Dual
I	L	yes	yes	$\|U\| = 1, \|X\| = 1$	$\{\wedge, \vee\}$-FO	I
II	NP-complete	yes	no	$\|U\| = 1, \|X\| \geq 2$	$\{\exists, \wedge, \vee\}$-FO	III
III	co-NP-complete	no	yes	$\|U\| \geq 2, \|X\| = 1$	$\{\forall, \vee, \wedge\}$-FO	II
IV	Pspace-complete	no	no	$\|U\| \geq 2, \|X\| \geq 2$	$\{\exists, \forall, \wedge, \vee\}$-FO	IV

Table 2: Reformulations of the tetrachotomy (U and X denote the subsets of the domain to which universal and existential variables relativise, respectively; the relativisation into a weaker logical fragment allows up to two constants).

surjective hyper-endomorphisms of a U-X-core in § 2.10; and, shows that it is unique up to isomorphism in § 2.11.

In § 3, we start by recalling Schaefer's theorem for Boolean CSP, and its analog for Boolean QCSP. In § 3.1, we deal with trivial fragments whose model checking problem is always in L. In 3.2, we turn our attention to fragments $\mathscr{L}$ whose complexity is trivial if the structure has a one element $\mathscr{L}$-core and is hard otherwise. In particular, we recall a Galois connection using hyper-endomorphisms to classify the fragment $\{\exists, \wedge, \vee\}$-FO following the guidelines given by Börner [Bör08] regarding Galois connections. In 3.3, we classify the fragments whose complexity can be deduced from the Boolean case.

In Section 41, we deal with $\{\exists, \forall, \wedge, \vee\}$-FO. In § 4.1, we recall the Galois connection using surjective hyper-endomorphisms. In § 4.2, we recall how the Boolean case can be classified using the lattice associated with this Galois connection. In general, the upper bound of our tetrachotomy is a direct consequence of the characterisation of U-X-core in terms of relativisation, and we only need to deal with the lower bounds which we do in full generality in § 4.3. In particular, in § 4.3.1, we characterise in some detail the DSM of a U-X-core, showing that it is of a very restricted form, which allows us to prove hardness in a generic way in subsequent sections. Finally, In § 4.4 we investigate the complexity of the meta-problem: given a finite structure $\mathscr{D}$, what is the complexity of evaluating positive equality-free sentences of FO over $\mathscr{D}$? We establish that the meta-problem is NP-hard, even for a fixed and finite signature.

The present paper represents the full version of the conference reports [Mar08, MM10, MM11] and a part of [MM12b]; [MM11] itself supersedes a series of papers begun with [MM09]. Unless otherwise stated, all results appear here for the first time (outside of conference publications).

2   Preliminaries

2.1   Basic Definitions

Unless otherwise stated, we shall work with finite relational structures that share the same finite relational signature σ. Let $\mathscr{D}$ be such a structure. We will denote its domain by D. We denote the size of such a set D by $|D|$. The *complement* $\overline{\mathscr{D}}$ of a structure $\mathscr{D}$ consists of relations that are exactly the set-theoretic complements of those in $\mathscr{D}$. I.e., for an a-ary R, $R^{\overline{\mathscr{D}}} := D^a \setminus R^{\mathscr{D}}$. For graphs this leads to a slightly non-standard notion of complement, as it includes self-loops.

A *homomorphism* (resp. *full homomorphism*) from a structure $\mathscr{D}$ to a structure $\mathscr{E}$ is a function $h : D \to E$ that preserves (resp. preserves fully) the relations of $\mathscr{D}$, i.e. for all a_i-ary relations R_i, and for all $x_1, \ldots, x_{a_i} \in D$, $R_i(x_1, \ldots, x_{a_i}) \in \mathscr{D}$ implies $R_i\big(h(x_1), \ldots, h(x_{a_i})\big) \in \mathscr{E}$ (resp. $R_i(x_1, \ldots, x_{a_i}) \in \mathscr{D}$ iff $R_i\big(h(x_1), \ldots, h(x_{a_i})\big) \in \mathscr{E}$). $\mathscr{D}$ and $\mathscr{E}$ are *homomorphically equivalent* if there are homomorphisms both

from $\mathscr{D}$ to $\mathscr{E}$ and from $\mathscr{E}$ to $\mathscr{D}$.

Let $\mathscr{L}$ be a fragment of FO. Let $\mathscr{D}$ be a fixed structure. The decision problem $\mathscr{L}(\mathscr{D})$ has:

- Input: a sentence φ of $\mathscr{L}$.

- Question: does $\mathscr{D} \models \varphi$?

2.2 Methodology

In this paper, we will be concerned with syntactic fragments $\mathscr{L}$ of FO defined by allowing or disallowing symbols from $\{\exists, \forall, \wedge, \vee, \neq, =, \neg\}$. Given any sentence φ in $\mathscr{L}$, we may compute in logarithmic space an equivalent sentence φ' in prenex normal form, with negation pushed inwards at the atomic level. Since we will not be concerned with complexities beneath L, we assume hereafter that all inputs are in this form.

In general Pspace membership of FO($\mathscr{D}$) follows by a simple evaluation procedure inward through the quantifiers. Similarly, the expression complexity of the existential fragment $\{\exists, \wedge, \vee, \neq, =\}$-FO is at most NP; and, that of its dual fragment $\{\forall, \vee, \wedge, =, \neq\}$-FO is at most co-NP (in both cases, we may even allow atomic negation) [Var82]. We introduce formally below this principle of duality.

Let $\mathscr{L}$ be a syntactic fragment of FO defined by allowing or disallowing symbols from $\{\exists, \forall, \wedge, \vee, \neq, =\}$. We denote by $\overline{\mathscr{L}}$ its dual fragment by de Morgan's law: $\wedge$ is dual to $\vee$, $\exists$ to $\forall$ and $=$ to $\neq$.

Proposition 1. *Let $\mathscr{L}$ be a syntactic fragment of* FO *defined by allowing or disallowing symbols from* $\{\exists, \forall, \wedge, \vee, \neq, =\}$. *The problem $\mathscr{L}(\mathscr{D})$ belongs to a complexity class* C *if, and only if, the problem $\overline{\mathscr{L}}(\overline{\mathscr{D}})$ belongs to the dual complexity class* co-C.

Proof. For any sentence φ in $\mathscr{L}$, we may rewrite its negation $\neg\varphi$ by pushing the negation inwards until all atoms appear negatively, denoting the sentence hence obtained by ψ (which is logically equivalent to $\neg\varphi$). Next, we replace every occurrence of a negated relational symbol $\neg R$ by R to obtain a sentence of $\overline{\mathscr{L}}$ which we denote by $\overline{\varphi}$. The following chain of equivalences holds

$$\mathscr{D} \models \varphi \iff \mathscr{D} \models \neg(\neg\varphi) \iff \mathscr{D} \models \neg(\psi) \iff \mathscr{D} \not\models \psi \iff \overline{\mathscr{D}} \not\models \overline{\varphi}.$$

Clearly, $\overline{\varphi}$ can be constructed in logspace from φ and the result follows. $\square$

We will use this principle of duality to only classify one fragment or its dual, for example we will study $\{\exists, \wedge\}$-FO and ignore its dual $\{\forall, \vee\}$-FO. We will also use this principle to classify the self-dual fragment $\{\exists, \forall, \wedge, \vee\}$-FO.

We assume at least one quantifier and one binary connective (weaker fragments being trivial). By the duality principle, we may consider only purely existential fragments, or fragments with both quantifiers. Regarding connectives, we have three possibilities: purely disjunctive fragments, purely conjunctive fragments and fragments with both connectives. Regarding equality and disequality, we should have the four possible subsets of $\{=, \neq\}$ but it will become clear that cases with both follow the same complexity delineation as the case with $\neq$ only. Moreover, for fragments with both quantifiers, we may use the duality principle between $\{\exists, \forall, \wedge\}$ and $\{\forall, \exists, \vee\}$ to simplify our task. This means that we would need to consider 3×3 positive existential fragments and 2×3 positive fragments with both quantifiers. Actually, we can decrease this last count by one, due to the duality between $\{\exists, \forall, \wedge, \vee, \neq\}$-FO and $\{\exists, \forall, \wedge, \vee, =\}$-FO. Regarding fragments with $\neg$, since we necessarily have both connectives and both quantifiers, we only have to consider two fragments: FO and $\{\exists, \forall, \wedge, \vee, \neg\}$-FO. However, we shall see that the complexity of FO agrees with that of $\{\exists, \forall, \wedge, \vee, \neq\}$-FO (and its dual $\{\exists, \forall, \wedge, \vee, =\}$-FO).

This makes a grand total of *15 fragments to classify*, which are listed below; the fragments marked with a $\star$ correspond to the CSP and QCSP and are still open. We will settle all other listed fragments.

The 15 relevant fragments can be organised broadly in the following four classes.

First Class This consists of the following trivial fragments: for such a fragment $\mathscr{L}$, the problem $\mathscr{L}(\mathscr{D})$ is trivial (in L) for any structure $\mathscr{D}$.

- $\{\exists, \vee\}$-FO   (see Proposition 30)

- $\{\exists, \vee, =\}$-FO   (see Proposition 30)

- $\{\exists, \vee, \neq\}$-FO   (see Proposition 30)

Second Class This consists of the following fragments which exhibit a simple dichotomy: for such a fragment $\mathscr{L}$, the problem $\mathscr{L}(\mathscr{D})$ is trivial (in L) when the $\mathscr{L}$-core (defined in the next section) of $\mathscr{D}$ has one element and hard otherwise (NP-complete for existential fragments, Pspace-complete for fragments that allow both quantifiers). For this class, tractability amounts to the relativisation of all quantifiers to some constant.

- $\{\exists, \wedge, \vee\}$-FO, $\{\exists, \wedge, \vee, =\}$-FO   (see Proposition 33.)

- $\{\exists, \wedge, \vee, \neq\}$-FO   (see Proposition 37.)

- $\{\exists, \forall, \wedge, \vee, \neq\}$-FO   (see Proposition 31.)

- $\{\exists, \forall, \wedge, \vee, \neg\}$-FO   (see Proposition 32.)

Third Class This exhibits more richness complexity-wise, tractability can not be explained simply by $\mathscr{L}$-core size and relativisation of quantifiers.

- $\{\exists, \wedge, \neq\}$-FO   (see Proposition 39.)

- $\{\exists, \forall, \wedge, \neq\}$-FO   (see Proposition 40.)

- ⋆ $\{\exists, \wedge\}$-FO, $\{\exists, \wedge, =\}$-FO

- ⋆ $\{\exists, \forall, \wedge\}$-FO, $\{\exists, \forall, \wedge, =\}$-FO

Fourth Class The last class consists of a single fragment and is rich complexity-wise, though we will see that a drop in complexity is always witnessed by relativisation of quantifiers.

- $\{\exists, \forall, \wedge, \vee\}$-FO   (see Theorem 41.)

2.3 $\mathscr{L}$-Containment and $\mathscr{L}$-Core

It is well known that conjunctive query containment is characterised by the presence of homomorphism between the corresponding canonical databases (this goes back to Chandra and Merlin [CM77], see also [GKL$^+$07, chapter 6]). For exactly the same reason, a similar result holds for $\{\exists, \wedge\}$-FO-containment. We state and prove this result for pedagogical reasons, before moving on to other fragments. The results in this section (§ 2.3) relating to existential fragments are essentially well known.

Let us fix some notation first. Given a primitive positive sentence φ in $\{\exists, \wedge\}$-FO, we denote by $\mathscr{D}_\varphi$ its *canonical database*, that is the structure with domain the variables of φ and whose tuples are precisely those that are atoms of φ. In the other direction, given a finite structure $\mathscr{A}$, we write $\varphi_\mathscr{A}$ for the so-called *canonical conjunctive query*[1] of $\mathscr{A}$, the quantifier-free formula that is the conjunction of

---

[1]Most authors consider the canonical query to be the sentence which is the existential quantification of $\varphi_\mathscr{A}$.

Fragment $\mathscr{L}$	$\mathscr{L}$-containment	$\mathscr{L}$-equivalence	$\mathscr{L}$-core
$\{\exists,\wedge\}$-FO $\{\exists,\wedge,=\}$-FO $\{\exists,\wedge,\vee\}$-FO $\{\exists,\wedge,\vee,\neq\}$-FO	homomorphism	homomorphic equivalence	(classical) core
$\{\exists,\forall,\wedge,\vee\}$-FO	surjective hyper-morphism	surjective hyper-morphism equivalence	U-X-core
$\{\exists,\forall,\wedge,\vee,\neg\}$-FO	Full surjective hyper-morphism	Full surjective hyper-morphism	quotient by $\sim$
contains $\{\exists,\wedge,\neq\}$-FO or contains $\{\forall,\vee,=\}$-FO	isomorphism	isomorphism	each structure

Table 3: The various notions of containment, equivalence and core for syntactic fragments of FO.

the positive facts of $\mathscr{A}$, where the variables $v_1,\ldots,v_{|A|}$ correspond to the elements $a_1,\ldots,a_{|A|}$ of $\mathscr{A}$. It is well known that $\mathscr{D}_\varphi$ is homomorphic to a structure $\mathscr{A}$ if, and only if, $\mathscr{A} \models \varphi$. Moreover, we now may define a winning strategy for $\exists$ in the Hintikka $(\mathscr{A},\varphi)$-game to be precisely the evaluation of the variables given by a homomorphism from $\mathscr{D}_\varphi$ to $\mathscr{A}$. Note also that $\mathscr{A}$ is isomorphic to the canonical database of $\exists v_1 \exists v_2 \ldots v_{|A|} \varphi_{\mathscr{A}}$.

Theorem 2. *Let $\mathscr{A}$ and $\mathscr{B}$ be two structures. The following are equivalent.*

(i) *For every sentence φ in $\{\exists,\wedge\}$-FO, if $\mathscr{A} \models \varphi$ then $\mathscr{B} \models \varphi$.*

(ii) *There exists a homomorphism from $\mathscr{A}$ to $\mathscr{B}$.*

(iii) *$\mathscr{B} \models \varphi_{\mathscr{A}}^{\{\exists,\wedge\}\text{-FO}}$ where $\varphi_{\mathscr{A}}^{\{\exists,\wedge\}\text{-FO}} := \exists v_1 \exists v_2 \ldots v_{|A|} \varphi_{\mathscr{A}}$.*

Proof. As we observed above, a homomorphism corresponds to a winning strategy in the $(\mathscr{A},\varphi)$-game and (ii) and (iii) are equivalent.

Clearly, (i) implies (iii) since $\mathscr{A} \models \exists v_1 \exists v_2 \ldots v_{|A|} \varphi_{\mathscr{A}}$.

We now prove that (ii) implies (i). Let h be a homomorphism from $\mathscr{A}$ to $\mathscr{B}$. If $\mathscr{A} \models \varphi$, then there is a homomorphism g from $\mathscr{D}_\varphi$ to $\mathscr{A}$. By composition, $h \circ g$ is a homomorphism from $\mathscr{D}_\varphi$ to $\mathscr{B}$. In other words, $h \circ g$ is a winning strategy for $\exists$ in the $(\mathscr{B},\varphi)$-game. $\square$

Definition 3. *Let $\mathscr{A}$ and $\mathscr{B}$ be two structures. We say that $\mathscr{A}$ is $\mathscr{L}$-contained in $\mathscr{B}$ if, and only if, for any φ in $\mathscr{L}$, $\mathscr{A} \models \varphi$ implies $\mathscr{B} \models \varphi$. We say that $\mathscr{A}$ and $\mathscr{B}$ are $\mathscr{L}$-equivalent if, and only if, for any φ in $\mathscr{L}$, $\mathscr{A} \models \varphi \Leftrightarrow \mathscr{B} \models \varphi$. If $\mathscr{B}$ is a minimal structure w.r.t. domain size such that $\mathscr{B}$ and $\mathscr{A}$ are $\mathscr{L}$-equivalent, then we say that $\mathscr{B}$ is an $\mathscr{L}$-core of $\mathscr{A}$.*

The $\{\exists,\wedge\}$-FO-core is unique up to isomorphism and is better known as *the core*. We proceed to characterise notions of containment, equivalence and core for other fragments of FO, which we will use to study the complexity of the associated model checking problems. The results of this section are summarised in Table 3.

Proposition 4. *Let $\mathscr{A}$ and $\mathscr{B}$ be two relational structures. The following are equivalent.*

8

(i) *There is a homomorphism from $\mathscr{A}$ to $\mathscr{B}$.*

(ii) *$\mathscr{A}$ is $\{\exists, \wedge\}$-FO-contained in $\mathscr{B}$.*

(iii) *$\mathscr{A}$ is $\{\exists, \wedge, =\}$-FO-contained in $\mathscr{B}$.*

(iv) *$\mathscr{A}$ is $\{\exists, \wedge, \vee\}$-FO-contained in $\mathscr{B}$.*

(v) *$\mathscr{A}$ is $\{\exists, \wedge, \vee, =\}$-FO-contained in $\mathscr{B}$.*

Proof. The equivalence of (i) and (ii) are stated in Theorem 2 and are equivalent to $\mathscr{B} \models \exists v_1 \exists v_2 \ldots v_{|A|} \varphi_{\mathscr{A}}$, a sentence of $\{\exists, \wedge\}$-FO. This takes care of the implications from (v), (iv) and (iii) to (i). Trivially (v) implies both (iv) and (iii).

It suffices to prove (i) implies (v). As in the proof of Theorem 2, it can be easily checked that a homomorphism can be applied to a winning strategy for $\exists$ in the $(\mathscr{A}, \varphi)$-game to obtain a winning strategy for $\exists$ in the $(\mathscr{B}, \varphi)$-game. To see this, write the quantifier-free part ψ of φ in conjunctive normal form as a disjunction of conjunction-of-positive-atoms ψ_i. We may even propagate equality out by substitution such that each ψ_i is equality-free (if some ψ_i contained no extensional symbol other than equality, the sentence φ would trivially holds on any structure as we only ever consider structures with at least one element). A winning strategy in the $(\mathscr{A}, \varphi)$-game corresponds to a homomorphism from some $\mathscr{D}_{\psi_i}$ to $\mathscr{A}$. By composition with the homomorphism from $\mathscr{A}$ to $\mathscr{B}$, we get a homomorphism from $\mathscr{D}_{\psi_i}$ to $\mathscr{B}$, *i.e.* a winning strategy in the $(\mathscr{B}, \varphi)$-game as required. □

Corollary 5. *Let $\mathscr{A}$ and $\mathscr{B}$ be two relational structures. The following are equivalent.*

(i) *$\mathscr{A}$ and $\mathscr{B}$ are homomorphically equivalent.*

(ii) *$\mathscr{A}$ and $\mathscr{B}$ have isomorphic cores.*

(iii) *$\mathscr{A}$ is $\{\exists, \wedge\}$-FO-equivalent to $\mathscr{B}$.*

(iv) *$\mathscr{A}$ is $\{\exists, \wedge, =\}$-FO-equivalent to $\mathscr{B}$.*

(v) *$\mathscr{A}$ is $\{\exists, \wedge, \vee\}$-FO-equivalent to $\mathscr{B}$.*

(vi) *$\mathscr{A}$ is $\{\exists, \wedge, \vee, =\}$-FO-equivalent to $\mathscr{B}$.*

We now move on to fragments containing $\{\exists, \wedge, \neq\}$.

Proposition 6. *Let $\mathscr{A}$ and $\mathscr{B}$ be two structures. The following are equivalent.*

(i) *For every sentence φ in $\{\exists, \wedge, \neq\}$-FO, if $\mathscr{A} \models \varphi$ then $\mathscr{B} \models \varphi$.*

(ii) *There exists an injective homomorphism from $\mathscr{A}$ to $\mathscr{B}$.*

(iii) *$\mathscr{B} \models \varphi_{\mathscr{A}}^{\{\exists, \wedge, \neq\}\text{-FO}}$ where $\varphi_{\mathscr{A}}^{\{\exists, \wedge, \neq\}\text{-FO}} := \exists v_1 \ldots v_{|A|} \varphi_{\mathscr{A}} \wedge \bigwedge_{1 \le i < j \le |A|} v_i \neq v_j$.*

Proof. Similar to Theorem 2. □

Corollary 7. *Let $\mathscr{L}$ be a fragment of FO such that $\mathscr{L}$ or its dual $\overline{\mathscr{L}}$ contains $\{\exists, \wedge, \neq\}$-FO. Let $\mathscr{A}$ and $\mathscr{B}$ be two structures. The following are equivalent.*

(i) *$\mathscr{A}$ and $\mathscr{B}$ are isomorphic.*

(ii) *$\mathscr{A}$ is $\mathscr{L}$-equivalent to $\mathscr{B}$.*

Proof. For the case when $\mathscr{L}$ contains $\{\exists, \wedge, \neq\}$-FO, the result follows from the previous proposition and the fact that we deal with finite structures only.

For the case when $\overline{\mathscr{L}}$ contains $\{\exists, \wedge, \neq\}$-FO, we apply the duality principle and the previous case, and equivalently $\overline{\mathscr{A}}$ and $\overline{\mathscr{B}}$ are isomorphic. This is in turn equivalent to $\mathscr{A}$ being isomorphic to $\mathscr{B}$. □

2.4 Hintikka Games

Before moving on to the equality-free fragments $\{\exists,\forall,\wedge,\vee\}$-FO and $\{\exists,\forall,\wedge,\vee,\neg\}$-FO, let us recall first basic definitions and notations regarding Hintikka Games. Let φ be a sentence of FO in prenex form with all negations pushed to the atomic level. A *strategy* for $\exists$ in the (Hintikka) $(\mathscr{A},\varphi)$-game is a set of mappings $\{\sigma_x : \text{`}\exists x\text{'} \in \varphi\}$ with one mapping σ_x for each existentially quantified variable x of φ. The mapping σ_x ranges over the domain A of $\mathscr{A}$; and, its domain is the set of functions from Y_x to A, where Y_x denotes the universally quantified variables of φ preceding x.

We say that $\{\sigma_x : \text{`}\exists x\text{'} \in \varphi\}$ is *winning* if for any assignment π of the universally quantified variables of φ to A, when each existentially quantified variable x is set according to σ_x applied to $\pi|_{Y_x}$, then the quantifier-free part ψ of φ is satisfied under this overall assignment h. When ψ is a conjunction of positive atoms, this amounts to h being a homomorphism from $\mathscr{D}_\psi$ to $\mathscr{A}$.

2.5 Hyper-morphisms

For the equality-free fragments $\{\exists,\forall,\wedge,\vee,\neg\}$-FO and $\{\exists,\forall,\wedge,\vee\}$-FO, the correct concept to transfer winning strategies involves unary hyper-operations, that is functions to the power-set.

A *hyper-operation* f from a set A to a set B is a function from A to the power-set of B. For a subset S of A, we will define its image $f(A)$ under the hyper-operation f as $\bigcup_{s \in S} f(s)$. When we wish to stress that an element may be sent to $\emptyset$, we speak of a *partial hyper-operation*; and otherwise we assume that f is *total*, that is for any a in A, $f(a) \neq \emptyset$. We say that f is *surjective* whenever $f(A) = B$. The *inverse* of a (total) hyper-operation f from A to B, denoted by f^{-1}, is the partial hyper-operation from B to A defined for any b in B as $f^{-1}(b) := \{a \in A \mid b \in f(a)\}$. We call an element of $f^{-1}(b)$ an *antecedent* of b under f. Let f be a hyper-operation from A to B and g a hyper-operation from B to C. The hyper-operation $g \circ f$ is defined naturally as $g \circ f(x) := g\big(f(x)\big)$ (recall that $f(x)$ is a set).

When f is a (total) surjective hyper-operation from A to A, we say that f is a *shop* of A. Note that the inverse of a shop is a shop and that the composition of two shops is a also a shop. Observing further that shop composition is associative and that the identity shop (which sends an element x of A to the singleton $\{x\}$) is the identity with respect to composition, we may consider the monoid generated by a set of shops. A shop f is a *sub-shop* of a shop g whenever, for every x in A, $f(x) \subseteq g(x)$. In our context, we will be interested in a particular monoid which will be closed further under sub-shops, a so-called *down-shop-monoid* (DSM). [2] We denote by $\langle F \rangle_{DSM}$ the DSM generated by a set F of shops.

Let f be a shop of A. When for a subset U of A we have $f(U) = A$, we say that f is *U-surjective*. Observing that the totality of f may be rephrased as $f^{-1}(A) = A$, we say more generally that f is *X-total* for a subset X of A whenever $f^{-1}(X) = A$. Note that for shops U-surjectivity and X-totality are dual to one another, that is the inverse of a U-surjective shop is an X-total shop with $X = U$ and vice versa. Somewhat abusing terminology, and when it does not cause confusion, we will drop the word surjective and by U- or U'-shop we will mean a U- or U'-surjective shop. Similarly, we will speak of an X- or X'-shop in the total case and of a U-X-shop in the case of a shop that is both U-surjective and X-total. Suitable compositions of U-shops and X-shops preserve these properties.

Lemma 8. *Let f and g be two shops.*

(i) *If f is a U-shop then $g \circ f$ is a U-shop.*

(ii) *If g is a X-shop then $g \circ f$ is a X-shop.*

(iii) *If both f is a U-shop and g is a X-shop then $g \circ f$ is a U-X-shop.*

---

[2]The "down" comes from *down-closure*, here under sub-shops; a nomenclature inherited from [Bör00].

(iv) *If both f and g are U-X-shops then $g \circ f$ is a U-X-shop.*

(v) *The iterate of a U-X-shop is a U-X-shop.*

Proof. We prove (i). Since $f(U) = A$, we have $g(f(U)) = g(A)$. By surjectivity of g, we know that $g(A) = A$. It follows that $g(f(U)) = A$ and we are done. (ii) is dual to (i), and (iii) follows directly from (i) and (ii). (iv) is a restriction of (iii) and is only stated here as we shall use it often. (v) follows by induction on the order of iteration using (iv). □

A *hyper-morphism f* from a structure $\mathscr{A}$ to a structure $\mathscr{B}$ is a hyper-operation from A to B that satisfies the following property.

- (**preserving**) if $R(a_1, \ldots, a_i) \in \mathscr{A}$ then $R(b_1, \ldots, b_i) \in \mathscr{B}$, for all $b_1 \in f(a_1), \ldots, b_i \in f(a_i)$.

When $\mathscr{A}$ and $\mathscr{B}$ are the same structure, we speak of a *hyper-endomorphism*. We say that f is *full* if moreover

- (**fullness**) $R(a_1, \ldots, a_i) \in \mathscr{A}$ iff $R(b_1, \ldots, b_i) \in \mathscr{B}$, for all $b_1 \in f(a_1), \ldots, b_i \in f(a_i)$.

Note that the inverse of a full surjective hyper-morphism is also a full surjective hyper-morphism.

2.6   Containment and Core for $\{\exists, \forall, \wedge, \vee, \neg\}$-FO

We now turn our attention to $\{\exists, \forall, \wedge, \vee, \neg\}$-FO. The proofs of the necessary characterisations are somewhat laboured, but will prepare us well for the forthcoming discussion on $\{\exists, \forall, \wedge, \vee\}$-FO.

Lemma 9. *Let $\mathscr{A}$ and $\mathscr{B}$ be two structures such that there is a full surjective hyper-morphism from $\mathscr{A}$ to $\mathscr{B}$. Then, for every sentence φ in $\{\exists, \forall, \wedge, \vee, \neg\}$-FO, if $\mathscr{A} \models \varphi$ then $\mathscr{B} \models \varphi$.*

Proof. Let h be a full surjective hyper-morphism from $\mathscr{A}$ to $\mathscr{B}$ and φ be a sentence of $\{\exists, \forall, \wedge, \vee, \neg\}$-FO such that $\mathscr{A} \models \varphi$. We fix an arbitrary linear order over A and write $\min h^{-1}(b)$ to denote the smallest antecedent of b in A under h.

Let $\{\sigma_x : `\exists x` \in \varphi\}$ be a winning strategy in the $(\mathscr{A}, \varphi)$-game. We construct a strategy $\{\sigma'_x : `\exists x` \in \varphi\}$ in the $(\mathscr{B}, \varphi)$-game as follows. Let $\pi_B : Y_x \to B$ be an assignment to the universal variables Y_x preceding an existential variable x in φ, we select for $\sigma'_x(\pi)$ an arbitrary element of $h(\sigma(\pi_A))$ where $\pi_A : Y_x \to A$ is an assignment such that for any universal variable y preceding x, we have $\pi_A(y) := \min h^{-1}(\pi_B(y))$. This strategy is well defined since h is surjective (which means that π_A is well defined) and total (which means that $h(\sigma(\pi_A)) \neq \emptyset$). Note moreover that using min in the definition of π_A means that a branch in the tree of the game on $\mathscr{B}$ will correspond to a branch in the tree of the game on $\mathscr{A}$. It remains to prove that $\{\sigma'_x : `\exists x` \in \varphi\}$ is winning. We will see that it follows from the fact that h is full and preserving.

We assume that negations have been pushed to the atomic level and write the quantifier-free part ψ of φ in disjunctive normal form as a disjunction of conjunctions-of-atoms ψ_i. If ψ_i has contradictory positive and negative atoms (as in $E(x,y) \wedge \neg E(x,y)$) then we may discard the sentence ψ_i as false. Moreover, for each pair of atoms $R(v_1, v_2, \ldots, v_r)$ and $\neg R(v_1, v_2, \ldots, v_r)$ (induced by the choice of a relational symbol R and the choice of r variables $v_1, v_2, \ldots, v_r$ occuring in ψ_i) such that neither is present in ψ_i, we may replace ψ_i by the logically equivalent $\big(\psi_i \wedge R(v_1, v_2, \ldots, v_r)\big) \vee \big(\psi_i \wedge \neg R(v_1, v_2, \ldots, v_r)\big)$. After this completion process, note that every conjunction of atoms ψ_i corresponds naturally to a structure $\mathscr{D}_{\psi_i}$ (take only the positive part of ψ_i which is now maximal).

Assume first that ψ is disjunction-free. The winning condition of the $(\mathscr{B}, \varphi)$-game can be recast as a full homomorphism from $\mathscr{D}_\psi$ to $\mathscr{B}$. Composing with h the full homomorphism from $\mathscr{D}_\psi$ to $\mathscr{A}$ (induced by the sequence of compatible assignments π_A to the universal variables and the strategy $\{\sigma_x : `\exists x` \in$

$\varphi\}$), we get a full hyper-morphism from $\mathscr{D}_\psi$ to $\mathscr{B}$. The map from the domain of $\mathscr{D}_\psi$ to $\mathscr{B}$ induced by the sequence of assignments π_B and the strategy $\{\sigma'_x : \text{`}\exists x\text{'} \in \varphi\}$ is a range restriction of this full hyper-morphism and is therefore a full homomorphism (we identify hyper-morphism to singletons with homomorphisms). In general when the quantifier-free part of φ has several disjuncts ψ_i, most likely after the completion process of the previous paragraph, the winning condition can be recast as a full homomorphism from some $\mathscr{D}_{\psi_i}$. The above argument applies and the result follows. $\square$

We shall see that the converse of Lemma 9 holds. Consequently, it turns out that containment and equivalence coincide for $\{\exists, \forall, \wedge, \vee, \neg\}$-FO, since the inverse of a full surjective hyper-morphism is a full surjective hyper-morphism.

For $\{\exists, \forall, \wedge, \vee, \neg\}$-FO, we define an equivalence relation $\sim$ over the structure elements in the spirit of the Leibnitz-rule for equality. For propositions P and Q, let $P \leftrightarrow Q$ be an abbreviation for $(P \wedge Q) \vee (\neg P \wedge \neg Q)$. For the sake of clarity, we deal with the case of digraphs first and write $x \sim y$ as an abbreviation for $\forall z(E(x,z) \leftrightarrow E(y,z)) \wedge (E(z,x) \leftrightarrow E(z,y))$. It is straightforward to verify that $\sim$ induces an equivalence relation over the vertices (which we denote also by $\sim$). In general, for each r-ary symbol R, let ψ_R stands for

$$\big(R(x,z_1,\ldots,z_{r-1}) \leftrightarrow R(y,z_1,\ldots,z_{r-1})\big) \wedge \big(R(z_1,x,z_2,\ldots,z_{r-1}) \leftrightarrow R(z_1,y,z_2,\ldots,z_{r-1})\big)$$
$$\wedge \ldots \wedge \big(R(z_1,z_2,\ldots,z_{r-1},x) \leftrightarrow R(z_1,z_2,\ldots,z_{r-1},y)\big).$$

We write $x \sim y$ for $\bigwedge_{R \in \sigma} \forall z_1, z_2, \ldots, z_{r-1} \psi_R$.

We write $\mathscr{A}/\sim$ for the quotient structure defined in the natural way. Note that there is a full surjective homomorphism from $\mathscr{A}$ to $\mathscr{A}/\sim$. As observed earlier, its inverse (viewing the homomorphism as an hyper-morphism) is a full surjective hyper-morphism from $\mathscr{A}/\sim$ to $\mathscr{A}$. Thus, it follows from Lemma 9 that $\mathscr{A}$ and $\mathscr{A}/\sim$ are $\{\exists, \forall, \wedge, \vee, \neg\}$-FO-equivalent.

Let $\varphi_{\mathscr{A}}^+$ denotes the (quantifier-free) canonical conjunctive query of $\mathscr{A}$ (denoted earlier as $\varphi_{\mathscr{A}}$) and $\varphi_{\mathscr{A}}^-$ denotes the similar sentence which lists the negative atoms of $\mathscr{A}$ instead of the positive atoms.

Proposition 10. *Let $\mathscr{A}$ and $\mathscr{B}$ be two structures. The following are equivalent.*

(i) *For every sentence φ in $\{\exists, \forall, \wedge, \vee, \neg\}$-FO, if $\mathscr{A} \models \varphi$ then $\mathscr{B} \models \varphi$.*

(ii) *There exists a full surjective hyper-morphism from $\mathscr{A}$ to $\mathscr{B}$.*

(iii) $\mathscr{B} \models \varphi_{\mathscr{A}}^{\{\exists, \forall, \wedge, \vee, \neg\}\text{-FO}}$ *where*

$$\varphi_{\mathscr{A}}^{\{\exists, \forall, \wedge, \vee, \neg\}\text{-FO}} := \exists v_1 \exists v_2 \ldots v_{|A|} \varphi_{\mathscr{A}}^+ \wedge \varphi_{\mathscr{A}}^- \wedge \forall w \bigvee_{1 \leq i \leq |A|} w \sim v_i.$$

(iv) *for every sentence φ in $\{\exists, \forall, \wedge, \vee, \neg\}$-FO, $\mathscr{A} \models \varphi$ iff $\mathscr{B} \models \varphi$.*

(v) $\mathscr{A}/\sim$ *and $\mathscr{B}/\sim$ are isomorphic.*

Proof. The implication (i) to (iii) is clear since by construction $\mathscr{A}$ models the canonical sentence $\varphi_{\mathscr{A}}^{\{\exists, \forall, \wedge, \vee, \neg\}\text{-FO}}$.

We prove that (iii) implies (ii). Assume that $\mathscr{B} \models \varphi_{\mathscr{A}}^{\{\exists, \forall, \wedge, \vee, \neg\}\text{-FO}}$. We construct a full and total surjective hyper-morphism h as follows. Let $b_1, b_2, \ldots, b_{|A|}$ be witnesses in B for $v_1, v_2, \ldots, v_{|A|}$. We set $h(a_i) \ni b_i$ for $1 \leq i \leq |A|$ (totality). For each b in B, we set the universal variable w to b and pick some j such that $w \sim v_j$ holds and set $h(a_j) \ni b$ (surjectivity). By construction, h is preserving and full.

The implication (ii) to (i) is proved as Lemma 9.

The equivalence of (i), (ii), (iii) with (iv) follows from our earlier observation that the inverse f^{-1} of a full surjective hyper-morphism f from $\mathscr{A}$ to $\mathscr{B}$ is a full surjective hyper-morphism from $\mathscr{B}$ to $\mathscr{A}$.

To see that (v) implies (ii), compose the quotient map from $\mathscr{A}$ to $\mathscr{A}/\sim$ (which is a full surjective homomorphism) with the inverse of the quotient map from $\mathscr{B}$ to $\mathscr{B}/\sim$ (which is a full surjective hyper-morphism).

For the direction (ii) to (v), the natural quotient $f/\sim$ of a full surjective hyper-morphism f from $\mathscr{A}$ to $\mathscr{B}$ is a full surjective homomorphism. Since we deal with finite structures, it is an isomorphism and we are done. $\square$

Note that no smaller structure can be $\{\exists, \forall, \wedge, \vee, \neg\}$-FO-equivalent to $\mathscr{A}' := \mathscr{A}/\sim$. Indeed, a full surjective hyper-morphism f from a smaller structure $\mathscr{B}$ to $\mathscr{A}'$ would have to satisfy $\{a'_1, a'_2\} \subseteq f(b)$ for some b in B and some distinct a'_1, a'_2 in A'. But this would imply that $a'_1 \sim a'_2$ which is not possible. Moreover, any structure that is $\{\exists, \forall, \wedge, \vee, \neg\}$-FO-equivalent and of the same size as $\mathscr{A}'$ will be isomorphic (a full surjective hyper-morphism must induce an isomorphism by triviality of $\sim$ over $\mathscr{A}'$). Thus, $\mathscr{A}/\sim$ is the (up to isomorphism unique) $\{\exists, \forall, \wedge, \vee, \neg\}$-FO-core of $\mathscr{A}$.

2.7   Containment for $\{\exists, \forall, \wedge, \vee\}$-FO

Lemma 11. *Let $\mathscr{A}$ and $\mathscr{B}$ be two structures such that there is a surjective hyper-morphism from $\mathscr{A}$ to $\mathscr{B}$. Then, for every sentence φ in $\{\exists, \forall, \wedge, \vee\}$-FO, if $\mathscr{A} \models \varphi$ then $\mathscr{B} \models \varphi$.*

Proof. The proof is exactly the same as that of Lemma 9, except that we no longer need to preserve atomic negation, and may drop the assumption of fullness. $\square$

We extend the notion of canonical conjunctive query of a structure $\mathscr{A}$. Given a tuple of (not necessarily distinct) elements $\mathbf{r} := (r_1, \ldots, r_l) \in A^l$, define the quantifier-free formula $\varphi_{\mathscr{A}(\mathbf{r})}(v_1, \ldots, v_l)$ to be the conjunction of the positive facts of $\mathbf{r}$, where the variables $v_1, \ldots, v_l$ correspond to the elements $r_1, \ldots, r_l$. That is, $R(v_{\lambda_1}, \ldots, v_{\lambda_i})$ appears as an atom in $\varphi_{\mathscr{A}(\mathbf{r})}$ iff $R(r_{\lambda_1}, \ldots, r_{\lambda_i})$ holds in $\mathscr{A}$. When $\mathbf{r}$ enumerates the elements of the structure $\mathscr{A}$, this definition coincides with the usual definition of canonical conjunctive query. Note also that in this case there is a full homomorphism from the canonical database $\mathscr{D}_{\varphi_{\mathscr{A}(\mathbf{r})}}$ to $\mathscr{A}$ given by the map $v_{\lambda_i} \mapsto r_i$.

Definition 12 ([MM12a]). *Let $\mathscr{A}$ be a structure and $m > 0$. Let $\mathbf{r}$ be an enumeration of the elements of $\mathscr{A}$.*

$$\theta_{\mathscr{A},m}^{\{\exists, \forall, \wedge, \vee\}\text{-FO}} := \exists v_1, \ldots, v_{|A|} \varphi_{\mathscr{A}(\mathbf{r})}(v_1, \ldots, v_{|A|}) \wedge \forall w_1, \ldots, w_m \bigvee_{\mathbf{t} \in A^m} \varphi_{\mathscr{A}(\mathbf{r}, \mathbf{t})}(\mathbf{v}, \mathbf{w}).$$

Observe that $\mathscr{A} \models \theta_{\mathscr{A},m}^{\{\exists, \forall, \wedge, \vee\}\text{-FO}}$. Indeed, we may take as witness for the variables $\mathbf{v}$ the corresponding enumeration $\mathbf{r}$ of the elements of $\mathscr{A}$; and, for any assignment $\mathbf{t} \in A^m$ to the universal variables $\mathbf{w}$, it is clear that $\mathscr{A} \models \varphi_{\mathscr{A}(\mathbf{r}, \mathbf{t})}(\mathbf{r}, \mathbf{t})$ holds.

Lemma 13. *Let $\mathscr{A}$ and $\mathscr{B}$ be two structures. If $\mathscr{B} \models \theta_{\mathscr{A}, |\mathscr{B}|}^{\{\exists, \forall, \wedge, \vee\}\text{-FO}}$ then there is a surjective hyper-morphism from $\mathscr{A}$ to $\mathscr{B}$.*

Proof. Let $\mathbf{b}' := b'_1, \ldots, b'_{|A|}$ be witnesses for $v_1, \ldots, v_{|A|}$. Assume that an enumeration $\mathbf{b} := b_1, b_2, \ldots, b_{|B|}$ of the elements of $\mathscr{B}$ is chosen for the universal variables $w_1, \ldots w_{|\mathscr{B}|}$. Let $\mathbf{t} \in A^m$ be the witness s.t. $\mathscr{B} \models \varphi_{\mathscr{A}(\mathbf{r})}(\mathbf{b}') \wedge \varphi_{\mathscr{A}(\mathbf{r}, \mathbf{t})}(\mathbf{b}', \mathbf{b})$.

Let f be the map from the domain of $\mathscr{A}$ to the power set of that of $\mathscr{B}$ which is the union of the following two partial hyper-operations h and g (i.e. $f(a_i) := h(a_i) \cup g(a_i)$ for any element a_i of $\mathscr{A}$), which guarantee totality and surjectivity, respectively.

- $h(a_i) := b'_i$ (totality.)

- $g(t_i) \ni b_i$ (surjectivity.)

It remains to show that f is preserving. This follows from $\mathscr{B} \models \varphi_{\mathscr{A}(\mathbf{r},\mathbf{t})}(\mathbf{b}',\mathbf{b})$.

Let R be a r-ary relational symbol such that $R(a_{i_1},\ldots,a_{i_r})$ holds in $\mathscr{A}$. Let $b''_{i_1} \in f(a_{i_1}),\ldots,b''_{i_r} \in f(a_r)$. We will show that $R(b''_{i_1},\ldots,b''_{i_r})$ holds in $\mathscr{B}$. Assume for clarity of the exposition and w.l.o.g. that from i_1 to i_k the image is set according to h and from i_{k+1} to i_r according to g: i.e. for $1 \leq j \leq k$, $h(a_{i_j}) = b'_{i_j} = b''_{i_j}$ and for $k+1 \leq j \leq r$, there is some l_j such that $t_{l_j} = a_{i_j}$ and $g(t_{l_j}) \ni b''_{i_j} = b_{l_j}$. By definition of $\mathscr{A}(\mathbf{r},\mathbf{t})$ the atom $R(v_{i_1},\ldots,v_{i_k},w_{l_{k+1}},\ldots,w_r)$ appears in $\varphi_{\mathscr{A}(\mathbf{r},\mathbf{t})}(\mathbf{v},\mathbf{w})$. It follows from $\mathscr{B} \models \varphi_{\mathscr{A}(\mathbf{r},\mathbf{t})}(\mathbf{b}',\mathbf{b})$ that $R(b''_{i_1},\ldots,b''_{i_r})$ holds in $\mathscr{B}$. $\qquad\square$

Theorem 14. *Let $\mathscr{A}$ and $\mathscr{B}$ be two structures. The following are equivalent.*

 (i) *For every sentence φ in $\{\exists,\forall,\wedge,\vee\}$-FO, if $\mathscr{A} \models \varphi$ then $\mathscr{B} \models \varphi$.*

 (ii) *There exists a surjective hyper-morphism from $\mathscr{A}$ to $\mathscr{B}$.*

 (iii) *$\mathscr{B} \models \theta_{\mathscr{A},|B|}^{\{\exists,\forall,\wedge,\vee\}\text{-FO}}$.*

Proof. By construction $\mathscr{A} \models \theta_{\mathscr{A},|B|}^{\{\exists,\forall,\wedge,\vee\}\text{-FO}}$, so (i) implies (iii). By Lemma 11, (ii) implies (i). By Lemma 13, (iii) implies (i). $\qquad\square$

2.8   A core for $\{\exists,\forall,\wedge,\vee\}$-FO

The property of a (classical) core can be rephrased in the logical context as the minimal $X = \widetilde{A} \subseteq A$ such that a primitive positive sentence φ is true on $\mathscr{A}$ iff it is true on $\mathscr{A}$ with the (existential) quantifiers relativised to $X = \widetilde{A}$. Let us say in this case that $\mathscr{A}$ has *X-relativisation* with respect to $\{\exists,\wedge\}$-FO.

Thus, the notion of a core can be recast in the context of $\{\exists,\wedge\}$-FO in a number of equivalent ways, as a minimal induced substructure $\widetilde{\mathscr{A}}$ of $\mathscr{A}$,

 (i) that satisfies the same $\{\exists,\wedge\}$-FO sentences;

 (ii) that is induced by minimal $X \subseteq A$ such that $\mathscr{A}$ has X-relativisation w.r.t. $\{\exists,\wedge\}$-FO; or,

 (iii) that is induced by minimal $X \subseteq A$ such that $\mathscr{A}$ has an endomorphism with image X.

We are looking for a useful characterisation of the analogous concept of core for $\{\exists,\forall,\wedge,\vee\}$-FO. As we now have both quantifiers, two sets U and X, one for each quantifier, will emerge naturally, hence we will call this core a *U-X-core*. As we shall see shortly, there are two equivalent ways of defining a U-X-core – one is logical, the other algebraic – as a minimal substructure $\widetilde{\mathscr{A}}$ of $\mathscr{A}$, induced by minimal $U,X \subseteq A$ such that:

 (ii) $\mathscr{A}$ has *$\forall U$-$\exists X$-relativisation* w.r.t. $\{\exists,\forall,\wedge,\vee\}$-FO; or,

 (iii) $\mathscr{A}$ has a *U-surjective X-total hyper-endomorphism*.

Recall that a surjective hyper-endomorphism f of $\mathscr{A}$ is *U-surjective* if $f(U) = A$ and *X-total* if $f^{-1}(X) = A$.

We will show that the sets U and X are unique up to isomorphism and that within a minimal induced substructure $\widetilde{\mathscr{A}}$, the sets U and X are uniquely determined. This will reconcile our definition of a U-X-core with the following natural definition, in which U and X are not explicit:

 (i) as a minimal induced substructure $\widetilde{\mathscr{A}}$ of $\mathscr{A}$ that satisfies the same sentences of $\{\exists,\forall,\wedge,\vee\}$-FO.

In our definition of $\{\exists,\forall,\wedge,\vee\}$-FO-core, we ask for a minimal structure, i.e. not necessarily an induced substructure. We shall see that it is equivalent to the above.

2.9   Relativisation

Given a formula φ, we denote by $\varphi_{[\forall u/\forall u \in U, \exists x/\exists x \in X]}$ the formula obtained from φ by relativising simultaneously every universal quantifier to U and every existential quantifier to X. When we only relativise universal quantifiers to U, we write $\varphi_{[\forall u/\forall u \in U]}$, and when we only relativise existential quantifiers to X, we write $\varphi_{[\exists x/\exists x \in X]}$.

Definition 15. *Let $\mathscr{A}$ be a finite structure over a set A, and U, X be two subsets of A. We say that $\mathscr{A}$ has $\forall U$-$\exists X$-relativisation if, for all sentences φ in $\{\exists, \forall, \wedge, \vee\}$-FO the following are equivalent*

(i) $\mathscr{A} \models \varphi$

(ii) $\mathscr{A} \models \varphi_{[\forall u/\forall u \in U]}$

(iii) $\mathscr{A} \models \varphi_{[\exists x/\exists x \in X]}$

(iv) $\mathscr{A} \models \varphi_{[\forall u/\forall u \in U, \exists x/\exists x \in X]}$

Lemma 16. *Let $\mathscr{A}$ be a finite structure over a set A, and U, X be two subsets of A. If $\mathscr{A}$ has a U-surjective X-total hyper-endomorphism then $\mathscr{A}$ has $\forall U$-$\exists X$-relativisation.*

Proof. Note that in Definition 15, we have $(iii) \Rightarrow (i) \Rightarrow (ii)$ and $(iii) \Rightarrow (iv) \Rightarrow (ii)$ trivially. It suffices to prove that $(ii) \Rightarrow (i)$ and $(i) \Rightarrow (iii)$ to complete the proof. To do so, we will consider the well known Hintikka game corresponding to Case (i), called the *unrelativised game* hereafter; and, the relativised Hintikka games corresponding to the relativised formulae from Cases (ii), (iii) and (iv) (the relativised game considered being clear from context).

Let h be a U-surjective X-total surjective hyper-endomorphism of $\mathscr{D}$. The proof follows the line of that of Lemma 11.

$((ii) \Rightarrow (i))$. Assume that we have a winning strategy in the universally relativised game. We produce a winning strategy in the unrelativised game using h. When taking the antecedent of a universal variable, we make sure to pick an antecedent in U which we can do by U-surjectivity of h. To be more precise, the linear order over $\mathscr{A}$ used in the proof of Lemma 11 starts with the elements of U.

$((i) \Rightarrow (iii))$. Assume that we have a winning strategy in the unrelativised game. We produce a winning strategy in the existentially relativised game using h. When taking the image of an existential variable, we no longer pick an arbitrary element but one in X, which we can do by X-totality of h. $\square$

Proposition 17. *The following are equivalent.*

(i) $\mathscr{A}$ *has $\forall U$-$\exists X$-relativisation.*

(ii) $\overline{\mathscr{A}}$ *has $\forall X$-$\exists U$-relativisation.*

Proof. It suffices to prove one implication. We prove (ii) implies (i). Let φ be a sentence of $\{\exists, \forall, \wedge, \vee\}$-FO. We use the duality principle and prove that $\mathscr{A} \models \varphi \iff \mathscr{A} \models \varphi_{[\forall u/\forall u \in U]}$. The other cases are similar and are omitted.

We follow the same notation as in Proposition 1: ψ is the sentence logically equivalent to $\neg\varphi$ with negation pushed at the atomic level, and $\overline{\varphi}$ is the sentence obtained from ψ by replacing every occurrence of a negative atom $\neg R$ by R. Recall the following chain of equivalence.

$$\mathscr{A} \models \varphi \iff \mathscr{A} \models \neg(\neg\varphi) \iff \mathscr{A} \models \neg(\psi) \iff \mathscr{A} \not\models \psi \iff \overline{\mathscr{A}} \not\models \overline{\varphi}.$$

By assumption $\overline{\mathscr{A}} \not\models \overline{\varphi} \iff \overline{\mathscr{A}} \not\models \overline{\varphi}_{[\exists u / \exists u \in U]}$. Using the above chain of equivalence backward and propagating the relativisation we obtain the following chain of equivalence.

$$\overline{\mathscr{A}} \not\models \overline{\varphi}_{[\exists u / \exists u \in U]} \iff \mathscr{A} \not\models \psi_{[\exists u / \exists u \in U]} \iff \mathscr{A} \models \neg(\psi_{[\exists u / \exists u \in U]})$$
$$\iff \mathscr{A} \models \neg(\neg \varphi_{[\forall u / \forall u \in U]}) \iff \mathscr{A} \models \varphi_{[\forall u / \forall u \in U]}.$$

$\square$

Lemma 18. *Let $\mathscr{A}$ be a finite structure over a set A, and U, X be two subsets of A. If $\mathscr{A}$ has $\forall U$-$\exists X$-relativisation then $\mathscr{A}$ has a U-surjective X-total hyper-endomorphism.*

Proof. Using the fact that the identity (defined as $i(x) := \{x\}$ for every x in $\mathscr{A}$) is a surjective hyper-endomorphism of $\mathscr{A}$ and applying Theorem 14, we derive that $\mathscr{A} \models \theta_{\mathscr{A}, |A|}^{\{\exists, \forall, \wedge, \vee\}\text{-FO}}$. By assumption, we may equivalently relativise only the existential quantifiers to X (Definition 15 (i) $\Rightarrow$ (iii)) and $\mathscr{A} \models \theta_{\mathscr{A}, |A|[\exists x / \exists x \in X]}^{\{\exists, \forall, \wedge, \vee\}\text{-FO}}$. Proceeding as in the proof of Lemma 11 but over this relativised sentence, we derive the existence of an X-total surjective hyper-operation g.

Using Proposition 17 and working over $\overline{\mathscr{A}}$, we derive similarly that $\overline{\mathscr{A}}$ has a U-total surjective hyper-operation. Let f be the inverse of this hyper-operation. Observe that it is a U-surjective hyper-operation.

By Lemma 8, the composition of these operations $g \circ f$ is a X-total U-surjective hyper-endomorphism as required. $\square$

Together, the two previous lemmata establish an algebraic characterisation of relativisation.

Theorem 19. *Let $\mathscr{A}$ be a finite structure over a set A, and U, X be two subsets of A. The following are equivalent.*

(i) *The structure $\mathscr{A}$ has $\forall U$-$\exists X$-relativisation.*

(ii) *The structure $\mathscr{A}$ has a X-total U-surjective hyper-endomorphism.*

Corollary 20. *Let $\mathscr{A}$ be a finite structure that has a U-surjective X-total hyper-endomorphism. Let $\widetilde{\mathscr{A}}$ be the substructure of $\mathscr{A}$ induced by $U \cup X$. The following holds.*

(i) *$\mathscr{A}$ and $\widetilde{\mathscr{A}}$ are $\{\exists, \forall, \wedge, \vee\}$-FO-equivalent.*

(ii) *$\widetilde{\mathscr{A}}$ has $\forall U$-$\exists X$-relativisation.*

Proof. Let f be the U-surjective X-total hyper-endomorphism of $\mathscr{A}$. Its range restriction g to $\widetilde{A} = U \cup X$ is a surjective hyper-morphism from $\mathscr{A}$ to $\widetilde{\mathscr{A}}$. The inverse g^{-1} of g is a surjective hyper-morphism from $\widetilde{\mathscr{A}}$ to $\mathscr{A}$, by X-totality of f. Appealing to Lemma 11 twice, once with g and once with g^{-1}, we obtain (i).

The restriction of g to $\widetilde{A}$ is a U-surjective X-total hyper-endomorphism of $\widetilde{\mathscr{A}}$, and (ii) follows from Lemma 16. $\square$

2.10 The U-X Core

Given a structure $\mathscr{D}$, we consider all minimal subsets X of D such that there is an X-total surjective hyper-endomorphism g of $\mathscr{D}$, and all minimal subsets U such that there is a U-surjective hyper-endomorphism f of $\mathscr{D}$. Such sets always exist as totality and surjectivity of surjective hyper-endomorphisms mean that in the worst case we may choose $U = X = D$. Since $g \circ f$ is a X-total U-surjective hyper-endomorphisms of $\mathscr{D}$ by Lemma 8, we may furthermore require that among all minimal sets satisfying the above, we choose a set U and a set X with $U \cap X$ maximal. Let $\widetilde{\mathscr{D}}$ be the substructure of $\mathscr{D}$ induced by $U \cup X$. We call $\widetilde{\mathscr{D}}$ a U-X-core of $\mathscr{D}$.

Remark 1. *Assume that there is an X_1-shop h_1 and an X_2-shop h_2 that preserves $\mathscr{D}$ such that $|X_1| > |X_2|$. We consider images of $h_1 \circ h_2$. For each element in X_2, pick a single element x_1' of X_1 in $h_1(X_2) \cap X_1$ such that $x_1' \in h_1(x_2)$. Let X_1' denote the set of picked elements. Since $|X_1| > |X_2|$ then $h_1 \circ h_2$ is an X_1'-shop that preserves $\mathscr{D}$ with $|X_1'| \le |X_2|$. Diagrammatically, this can be written as,*

$$D \xrightarrow{h_2} X_2 \xrightarrow{h_1} X_1' \subseteq h_1(X_2) \cap X_1 \subseteq X_1 \subseteq h_1 \circ h_2(D).$$

This means that we may look for an X-shop where the set X is minimal with respect to inclusion, or equivalently, for a set with minimal size $|X|$. So, in order to find an X-shop with a minimal set $|X|$, we may proceed greedily, removing elements from D while we have an X-shop until we obtain a set X such that there is no X'-shop for $X' \subsetneq X$. The dual argument applies to U-shops, and consequently to U-X-shops.

This further explains why minimising U and X, and then maximising their intersection, necessarily leads to a minimal $\widetilde{D} := U \cup X$ also. Because, would we find $U' \cup X'$ of smaller size, we might look within U' and X' for potentially smaller sets of cardinality $|U|$ and $|X|$, thus contradicting minimality.

Note that the sets U and X are not necessarily unique. However, as we shall see later *the U-X-core is unique up to isomorphism* (see Theorem 27). Moreover, within $\widetilde{\mathscr{D}}$, the sets U and X are uniquely determined. We delay until later the proof of this second result (see Theorem 50).

2.11   Uniqueness of the *U-X*-core

Throughout this section, let $\mathscr{D}$ be a finite structure and $\mathscr{M}$ its associated DSM; i.e. $\mathscr{M}$ is the set of surjective hyper-endomorphisms of $\mathscr{D}$. Let U and X be subsets of D such that the substructure $\widetilde{D}$ of $\mathscr{D}$ induced by $\widetilde{D} = U \cup X$ is a U-X-core of $\mathscr{D}$. We will progress through various lemmata and eventually derive the existence of a canonical U-X-shop in $\mathscr{M}$ which will be used to prove that the U-X-core is unique up to isomorphism. Uniqueness of the U-X-core has no real bearing on our classification program but the canonical shop will allow us to characterise all other shops in $\mathscr{M}$, which will be instrumental in the hardness proofs for $\{\exists, \forall, \wedge, \vee\}$-FO$(\mathscr{D})$.

Lemma 21. *Let f be a shop in $\mathscr{M}$. For any element z in D, $f(z)$ contains at most one element of the set U, that is $|f(z) \cap U| \le 1$.*

Proof. Assume for contradiction that there is some z and some distinct elements u_1 and u_2 of U such that $f(z) \supseteq \{u_1, u_2\}$. Let $z_3, z_4, \ldots$ be any choice of antecedents under f of the remaining elements $u_3, u_4, \ldots$ of U (recall that f is surjective). By assumption the monoid $\mathscr{M}$ contains a U-shop g. Hence, $g \circ f$ would be a U'-shop with $U' = \{z, z_3, z_4, \ldots\}$ since $f(U') \subseteq U$ and $g(U) = D$. We get a contradiction as $|U'| < |U|$. $\qquad\square$

Lemma 22. *Let f be a U-shop in $\mathscr{M}$. There exists a permutation α of U such that: for any u in U,*

(i) *$f(u) \cap U = \{\alpha(u)\}$; and,*

(ii) *$f^{-1}(u) \cap U = \{\alpha^{-1}(u)\}$.*

Proof. It follows from Lemma 21 that for any u in U, $|f(u) \cap U| \le 1$. Since f is a U-shop, every element in D has an antecedent in U under f and thus in particular for any u in U, $|f^{-1}(u) \cap U| \ge 1$. Note that if some element of U had no image in U then as U is finite, we would have an element of U with two distinct images in U. Hence, for any u in U, $|f(u) \cap U| = 1$ and the result follows. $\qquad\square$

The dual statements concerning X-shops hold.

Lemma 23. *Let f be a shop in $\mathcal{M}$. for any element z in D, $f^{-1}(z)$ contains at most one element of the set X, that is $|f^{-1}(z)\cap X| \leq 1$.*

Proof. By duality from Lemma 21. □

Lemma 24. *Let f be an X-shop in $\mathcal{M}$. There exists a permutation β of X such that: for any x in X,*

(i) $f(x)\cap X = \{\beta(x)\}$; *and,*

(ii) $f^{-1}(x)\cap X = \{\beta^{-1}(x)\}$.

Proof. By duality from Lemma 22. □

Lemma 25. *Let f be a shop in $\mathcal{M}$. If f is a U-X-shop then $f(X)\cap(U\setminus X) = \emptyset$.*

Proof. Assume for contradiction that for some $x_1 \in X$ and some $u_1 \in U\setminus X$, we have $u_1 \in f(x_1)$. Since f is an X-shop, every element is an antecedent under f of some element in X, in particular every element $x_2, x_3, \ldots \in X$ (different from x_1) has a unique image $x'_2, x'_3, \ldots \in X$ (see Lemma 24). Some element of X, say x_i does not occur in these images. Necessarily, x_1 reaches x_i. Note that x_i can not also belong to U as otherwise, x_i and u_1, two distinct elements of U, would be reached by x_1, contradicting Lemma 21. Thus, we must have that x_i belongs to $X\setminus U$. Let $U' := U$ and $X' := X\setminus\{x_i\}\cup\{u_1\}$. Note that $f^2 := f\circ f$, the second iterate of f, is a U'-X'-shop with $|U'| = |U|$, $|X'| = |X|$ and $|U'\cap X'| < |U\cap X|$. This contradicts our hypothesis on U and X. □

Proposition 26. *Let $\mathcal{M}$ be a DSM over a set D and U and X be minimal subsets of D such that: there is a U-shop in $\mathcal{M}$; there is an X-shop in $\mathcal{M}$ and $U\cup X$ is minimal. Then, there is a U-X-shop h in $\mathcal{M}$ that has the following properties:*

(i) *for any y in $U\cap X$, $h(y)\cap(U\cup X) = \{y\}$;*

(ii) *for any x in $X\setminus U$, $h(x)\cap(U\cup X) = \{x\}$;*

(iii) *for any u in $U\setminus X$, $h(u)\cap(U\cup X) = \{u\}\cup X_u$, where $X_u \subseteq X\setminus U$; and,*

(iv) *$h(U\setminus X)\cap X = \bigcup_{u\in U\setminus X} X_u = X\setminus U$.*

Proof. By assumption, $\mathcal{M}$ contains a U-X-shop f. Let α and β be permutations of U and X, respectively, as in Lemmata 22 and 24. Let r be the least common multiple of the order of the permutations β and α. We set h to be the rth iterate of f and we now know that $h(z) \ni z$ for any element z and that h is a U-X-shop by 8. Let y in $U\cap X$, we know that $h(y) \ni y$. We can not have another element from $U\cup X$ in $h(y)$ by Lemmata 21 and 24. This proves (i). Let x in $X\setminus U$, we know that $h(x) \ni x$. We can not have an element from X distinct from x in $h(x)$ by Lemma 24 and we can not have an element from $U\setminus X$ in $h(x)$ by Lemma 25. This proves (ii). Let u in $U\setminus X$, we know that $h(u) \ni u$. We can not have an element from U distinct from u in $h(u)$ by Lemma 22. We may have however some elements from $X\setminus U$ in $h(u)$. Thus, there is a set $\emptyset \subseteq X_u \subseteq X\setminus U$ such that $h(u)\cap(U\cup X) = \{u\}\cup X_u$. This proves (iii). By construction h is a U-shop and every element must have an antecedent in U under h. Since by the first three points, elements from $X\setminus U$ can only be reached from elements in $U\setminus X$, the last point (iv) follows. □

Theorem 27. *The U-X-core is unique up to isomorphism.*

Proof. Let h_1 be a U_1-X_1-shop with minimal $|U_1|$, $|X_1|$ and $|U_1 \cup X_1|$ and let h_2 be a U_2-X_2-shop with minimal $|U_2|$, $|X_2|$ and $|U_2 \cup X_2|$. Hence, $h_1 \circ h_2$ is a $h_1(X_2) \cap X_1$-shop with $|h_1(X_2)| \leq |X_1|$. By minimality of X_1, $|h_1(X_2)| = |X_1|$, and the restriction of h_1 to domain X_2 and codomain X_1 induces a surjective homomorphism from the substructure induced by X_2 to the substructure induced by X_1. Similarly h_2 induces a surjective homomorphism in the other direction. As we work with finite structures, h_1 induces an isomorphism i from the substructure induced by X_1 to the substructure induced by X_2. By duality, we also get that h_1 induces an isomorphism i' from the substructure induced by U_1 to the substructure induced by U_2. By construction, i and i' agree on $U_1 \cap X_1$ (necessarily to $U_2 \cap X_2$) and the result follows.   □

3   Complexity classification

We recall below some well known results concerning the complexity of Boolean CSP and QCSP which we will need later. For definitions and further details regarding the proof, the reader may consult the nice survey by Chen [Che09].

Theorem 28 ([Sch78])**.** *Let $\mathscr{D}$ be a Boolean structure. Then $CSP(\mathscr{D})$ (equivalently, $\{\exists, \wedge\}$-FO$(\mathscr{D})$) is in* P *if, and only if, all relations of $\mathscr{D}$ are simultaneously 0-valid, 1-valid, Horn, dual-Horn, bijunctive or affine, and otherwise it is* NP-*complete.*

A similar result holds when universal quantifiers are added to the mix. (it was sketched in the presence of constants [Sch78], then proved in the absence of constants in [CKS01] and [Dal97]).

Theorem 29. *Let $\mathscr{D}$ be a Boolean structure. Then $QCSP(\mathscr{D})$ (equivalently, $\{\exists, \forall, \wedge\}$-FO$(\mathscr{D})$) is in* P *if, and only if, all relations of $\mathscr{D}$ are simultaneously Horn, dual-Horn, bijunctive or affine, and otherwise it is* Pspace-*complete.*

Example 1. *The canonical example of a relation that does not fall in any of the tractable cases is* NAE $:= \{0,1\}^3 \setminus \{(0,0,0),(1,1,1)\}$. *Let $\mathscr{B}_{NAE}$ be the Boolean structure with this relation. It follows from the above theorems that $CSP(\mathscr{B}_{NAE})$ is* NP-*complete and that $QCSP(\mathscr{B}_{NAE})$ is* Pspace-*complete.*

Example 2. *For larger domains, though the classification remains open, the canonical hard problem is induced by the relation $\neq$. Let $\mathscr{K}_n$ denote the clique of size n (we view an undirected graph as a structure with a single binary predicate E that is symmetric). For $n \geq 3$, $CSP(\mathscr{K}_n)$ is a reformulation of the n-colourability problem and is* NP-*complete. It is also known that for $n \geq 3$ $QCSP(\mathscr{K}_n)$ is* Pspace-*complete [BBC$^+$09].*

3.1   First Class

Proposition 30.   (i) *When $\mathscr{D}$ has a single element, the model checking problem for* FO *is in* L.

  (ii) *The model checking problem $\{\exists, \vee, \neq, =\}$-FO is in* L.

Proof.   (i) In the case where $|D| = 1$, every relation is either empty or contains all tuples (one tuple), and the quantifiers $\exists$ and $\forall$ are semantically equivalent. Hence, the problem translates to the Boolean Sentence Value Problem (under the substitution of 0 and 1 for the empty and non-empty relations, respectively), known to be in L [Lyn77].

  (ii) We may assume by the previous point that $|D| > 1$. We only need to check if one of the atoms that occurs as a disjunct in the input sentence holds in $\mathscr{D}$. Since $|D| > 1$, a sentence with an atom like $x = y$ or $x \neq y$ is always true in $\mathscr{D}$. For sentences of $\{\exists, \vee\}$-FO, the atoms may have implicit equality as in $R(x,x,y)$ for a ternary predicate R: in any case, each atom may be checked in constant time since $\mathscr{D}$ is a fixed structure, resulting in overall logspace complexity.   □

     □

3.2 Second Class

We now move on to the largest fragments we will need to consider, which turn out to exhibit trivial dichotomies.

Proposition 31. *In full generality, the class of problems $\{\exists, \forall, \wedge, \vee, \neq\}$-FO$(\mathscr{D})$ exhibits dichotomy: if $|D| = 1$ then the problem is in* L*, otherwise it is* Pspace*-complete. Consequently, the fragment extended with $=$ follows the same dichotomy.*

Proof. When $|D| \geq 2$, Pspace-hardness may be proved using no extensional relation of D other than $\neq$. The formula $\varphi_{\mathscr{K}_{|D|}}(x,y) := (x \neq y)$ simulates the edge relation of the clique $\mathscr{K}_{|\mathscr{D}|}$ and the problem $\{\exists, \forall, \wedge\}$-FO$(\mathscr{K}_n)$ better known as QCSP$(\mathscr{K}_n)$ is Pspace-complete for $n \geq 3$ [BBC$^+$09]. For n = 2, we use a reduction from the problem QCSP$(\mathscr{B}_{\mathrm{NAE}})$ to prove that $\{\exists, \forall, \wedge, \vee\}$-FO$(\mathscr{K}_2)$ is Pspace-complete. Let φ be an input for QCSP$(\mathscr{B}_{\mathrm{NAE}})$. Let φ' be built from φ by substituting all instances of NAE(x,y,z) by $E(x,y) \vee E(y,z) \vee E(x,z)$. It is easy to see that $\mathscr{B}_{\mathrm{NAE}} \models \varphi$ iff $\mathscr{K}_2 \models \varphi'$, and the result follows.

Note that we have not used $=$ in our hardness proof; and, in the case $|D| = 1$, we may allow $=$ without affecting tractability (triviality). Thus, the fragment extended with $=$ follows the same delineation. $\square$

Proposition 32. *In full generality, the class of problems $\{\exists, \forall, \wedge, \vee, \neg\}$-FO$(\mathscr{D})$ exhibits dichotomy: if all relations of $\mathscr{D}$ are* trivial *(either empty or contain all tuples) then the problem is in* L*, otherwise it is* Pspace*-complete.*

Proof. If all relations are trivial then $\sim$ has a single equivalence class. Thus, $\mathscr{D}/\!\!\sim$ has a single element. By $\{\exists, \forall, \wedge, \vee, \neg\}$-FO-equivalence of $\mathscr{D}$ and $\mathscr{D}/\!\!\sim$ (see Proposition 10), it suffices to check whether an input φ in $\{\exists, \forall, \wedge, \vee, \neg\}$-FO holds in $\mathscr{D}/\!\!\sim$. Since the latter has a single element, the problem is in L.

Otherwise, the equivalence relation $\sim$ has at least 2 equivalence classes since $\mathscr{D}$ is non trivial. We may now follow the same proof as in Proposition 31, using the negation of $\sim$ in lieu of $\neq$, and Pspace-hardness follows. $\square$

We proceed with the last two fragments of the second class.

Proposition 33. *In full generality, the class of problems $\{\exists, \wedge, \vee\}$-FO$(\mathscr{D})$ exhibits dichotomy: if the core of $\mathscr{D}$ has one element then the problem is in* L*, otherwise it is* NP*-complete. As a corollary, the class of problems $\{\exists, \wedge, \vee, =\}$-FO$(\mathscr{D})$ exhibits the same dichotomy.*

In our preliminary work [Mar08, Mar06], the proof of the above is combinatorial and appeals to Hell and Nešetřil's dichotomy theorem for undirected graphs [HN90]. An alternative proof of this result also appeared in [BHR09] (and to a lesser extent in [HR09]). We will give here an algebraic proof which uses a variant of the Galois connection $\mathsf{Inv} - \mathsf{End}$ due to Krasner [Kra38] for $\{\exists, \wedge, \vee, =\}$-FO. The variant for $\{\exists, \wedge, \vee\}$-FO involves hyper-endomorphisms rather than endomorphisms because of the absence of equality. A hyper-endomorphism of $\mathscr{B}$ is a function from B to the power-set of B that is total and preserving (see Definition 2.5). Our purpose is to provide both a self-contained proof and a gentle introduction to the techniques we shall use for the fragment $\{\exists, \forall, \wedge, \vee\}$-FO.

For a set F of hyper-endomorphisms on the finite domain B, let $\mathsf{Inv}(F)$ be the set of relations on B of which each f in F is an hyper-endomorphism (when these relations are viewed as a structure over B). We say that $S \in \mathsf{Inv}(F)$ is invariant or *preserved* by (the hyper-endomorphisms in) F. Let $\mathsf{hE}(\mathscr{B})$ be the set of hyper-endomorphisms of $\mathscr{B}$. Let $\langle\mathscr{B}\rangle_{\{\exists, \wedge, \vee\}\text{-FO}}$ be the set of relations that may be defined on $\mathscr{B}$ in $\{\exists, \wedge, \vee\}$-FO.

Lemma 34. *Let $\mathbf{r} := (r_1, \ldots, r_k)$ be a k-tuple of elements of $\mathscr{B}$. There exists a formula $\theta_{\mathbf{r}}^{\{\exists, \wedge, \vee\}\text{-FO}}(u_1, \ldots, u_k) \in \{\exists, \wedge, \vee\}$-FO such that the following are equivalent.*

(i) $(\mathscr{B}, r'_1, \ldots, r'_k) \models \theta_{\mathbf{r}}^{\{\exists, \wedge, \vee\}\text{-FO}}(u_1, \ldots, u_k)$.

(ii) *There is a hyper-endomorphism from $(\mathscr{B}, r_1, \ldots, r_k)$ to $(\mathscr{B}, r'_1, \ldots, r'_k)$.*

Proof. let $\mathbf{s} := (b_1, \ldots, b_{|B|})$ an enumeration of the elements of $\mathscr{B}$ and $\varphi_{\mathscr{B}(\mathbf{r},\mathbf{s})}(v_1, \ldots, v_{|B|})$ be the associated conjunction of positive facts. Set

$$\theta_{\mathbf{r}}^{\{\exists, \wedge, \vee\}\text{-FO}}(u_1, \ldots, u_k) := \exists v_1, \ldots, v_{|B|} \; \varphi_{\mathscr{B}(\mathbf{r},\mathbf{s})}(v_1, \ldots, v_{|B|}).$$

The forward direction is clear as the witness $s'_1, \ldots, s'_{|B|}$ for $v_1, \ldots, v_{|B|}$ provides a hyper-endomorphism f defined as $f(b_i) \ni s'_i$ and $f(r_i) \ni r'_i$.

For the backwards direction, one may build an endomorphism from $(\mathscr{B}, r_1, \ldots, r_k)$ to $(\mathscr{B}, r'_1, \ldots, r'_k)$ from the given hyper-endomorphism. The result follows from the implication from (i) to (iv) of Proposition 4. $\qquad\square$

Theorem 35. *For a finite structure $\mathscr{B}$ we have $\langle\mathscr{B}\rangle_{\{\exists, \wedge, \vee\}\text{-FO}} = \mathsf{Inv}(\mathsf{hE}(\mathscr{B}))$.*

Proof. Let $\varphi(\mathbf{v})$ be a formula of $\{\exists, \wedge, \vee\}$-FO with free variables $\mathbf{v}$. We denote also by $\varphi(\mathbf{v})$ the relation induced over $\mathscr{B}$.

1. $\varphi(\mathbf{v}) \in \langle\mathscr{B}\rangle_{\{\exists, \wedge, \vee\}\text{-FO}} \Rightarrow \varphi(\mathbf{v}) \in \mathsf{Inv}(\mathsf{hE}(\mathscr{B}))$. This is proved by induction on the complexity of $\varphi(\mathbf{v})$.

   (Base Cases.) When $\varphi(\mathbf{v}) := R(\mathbf{v})$, the variables $\mathbf{v}$ may appear multiply in R and in any order. Thus R is an instance of an extensional relation under substitution and permutation of positions. The result follows directly from the definition of hyper-endomorphisms.

   (Inductive Step.) There are three subcases. We progress through them in a workmanlike fashion. Take $f \in \mathsf{hE}(\mathscr{B})$.

   (a) $\varphi(\mathbf{v}) := \psi(\mathbf{v}) \wedge \psi'(\mathbf{v})$. Let $\mathbf{v} := (v_1, \ldots, v_l)$. Suppose $\mathscr{B} \models \varphi(x_1, \ldots, x_l)$; then both $\mathscr{B} \models \psi(x_1, \ldots, x_l)$ and $\mathscr{B} \models \psi'(x_1, \ldots, x_l)$. By Inductive Hypothesis (IH), for any $y_1 \in f(x_1), \ldots, y_l \in f(x_l)$, both $\mathscr{B} \models \psi(y_1, \ldots, y_l)$ and $\mathscr{B} \models \psi'(y_1, \ldots, y_l)$, whence $\mathscr{B} \models \varphi(y_1, \ldots, y_l)$.

   (b) $\varphi(\mathbf{v}) := \psi(\mathbf{v}) \vee \psi'(\mathbf{v})$. Let $\mathbf{v} := (v_1, \ldots, v_l)$. Suppose $\mathscr{B} \models \varphi(x_1, \ldots, x_l)$; then one of $\mathscr{B} \models \psi(x_1, \ldots, x_l)$ or $\mathscr{B} \models \psi'(x_1, \ldots, x_l)$; w.l.o.g. the former. By IH, for any $y_1 \in f(x_1), \ldots, y_l \in f(x_l)$, $\mathscr{B} \models \psi(y_1, \ldots, y_l)$, whence $\mathscr{B} \models \varphi(y_1, \ldots, y_l)$.

   (c) $\varphi(\mathbf{v}) := \exists w \; \psi(\mathbf{v}, w)$. Let $\mathbf{v} := (v_1, \ldots, v_l)$. Suppose $\mathscr{B} \models \exists w \; \psi(x_1, \ldots, x_l, w)$; then for some x', $\mathscr{B} \models \psi(x_1, \ldots, x_l, x')$. By IH, for any $y_1 \in f(x_1), \ldots, y_l \in f(x_l), y' \in f(x')$, $\mathscr{B} \models \psi(y_1, \ldots, y_l, y')$, whereupon $\mathscr{B} \models \exists w \; \psi(y_1, \ldots, y_l, w)$.

2. $S \in \mathsf{Inv}(\mathsf{hE}(\mathscr{B})) \Rightarrow S \in \langle\mathscr{B}\rangle_{\{\exists, \wedge, \vee\}\text{-FO}}$. Consider the k-ary relation $S \in \mathsf{Inv}(\mathsf{hE}(\mathscr{B}))$. Let $\mathbf{r}_1, \ldots, \mathbf{r}_m$ be the tuples of S. Set $\theta_S^{\{\exists, \wedge, \vee\}\text{-FO}}(u_1, \ldots, u_k)$ to be the following formula of $\{\exists, \wedge, \vee\}$-FO:

$$\theta_{\mathbf{r}_1}^{\{\exists, \wedge, \vee\}\text{-FO}}(u_1, \ldots, u_k) \vee \ldots \vee \theta_{\mathbf{r}_m}^{\{\exists, \wedge, \vee\}\text{-FO}}(u_1, \ldots, u_k).$$

   For $\mathbf{r}_i := (r_{i1}, \ldots, r_{ik})$, note that $(\mathscr{B}, r_{i1}, \ldots, r_{ik}) \models \theta_{\mathbf{r}_i}^{\{\exists, \wedge, \vee\}\text{-FO}}(u_1, \ldots, u_k)$. That $\theta_S(u_1, \ldots, u_k) = S$ now follows from Part (ii) of Lemma 34, since $S \in \mathsf{Inv}(\mathsf{hE}(\mathscr{B}))$.

$\qquad\square$

Note that in [HR09] it is erroneously claimed that endomorphisms (not hyper-endomorphisms) are the correct algebraic object for the fragment $\{\exists, \wedge, \vee\}$-FO – this is not correct and only holds for the richer fragment $\{\exists, \wedge, \vee, =\}$-FO.

Corollary 36. *Let $\mathscr{B}$ and $\mathscr{B}'$ be finite structures over the same domain B. If $\mathsf{hE}(\mathscr{B}) \subseteq \mathsf{hE}(\mathscr{B}')$ then $\{\exists, \wedge, \vee\}$-FO$(\mathscr{B}') \leq_{\mathrm{L}} \{\exists, \wedge, \vee\}$-FO$(\mathscr{B})$.*

Proof. If $\mathsf{hE}(\mathscr{B}) \subseteq \mathsf{hE}(\mathscr{B}')$, then $\mathsf{Inv}(\mathsf{hE}(\mathscr{B}')) \subseteq \mathsf{Inv}(\mathsf{hE}(\mathscr{B}))$. From Theorem 35, it follows that $\langle \mathscr{B}' \rangle_{\{\exists, \wedge, \vee\}\text{-FO}}$ $\subseteq \langle \mathscr{B} \rangle_{\{\exists, \wedge, \vee\}\text{-FO}}$. Recalling that $\mathscr{B}'$ contains only a finite number of extensional relations, we may therefore effect a Logspace reduction from $\{\exists, \wedge, \vee\}$-FO$(\mathscr{B}')$ to $\{\exists, \wedge, \vee\}$-FO$(\mathscr{B})$ by straightforward substitution of predicates. $\square$

of Proposition 33. By Proposition 5, we may assume w.l.o.g. that $\mathscr{D}$ is a core. This means that every hyper-endomorphism of $\mathscr{D}$ is in fact an automorphism – we identify hyper-endomorphisms whose range are singletons with automorphisms – and thus $\mathsf{hE}(\mathscr{D})$ is a subset of S_n where $n = |D|$. If D has one element, then the problem is trival. If D has two elements, then $\mathsf{hE}(\mathscr{D}) \subseteq \mathsf{hE}(\mathscr{B}_{\mathrm{NAE}}) = S_2$. By Lemma 36, it follows that $\{\exists, \wedge, \vee\}$-FO$(\mathscr{B}_{\mathrm{NAE}}) \leq_{\mathrm{L}} \{\exists, \wedge, \vee\}$-FO$(\mathscr{D})$. Since the former is a generalisation of the NP-complete CSP$(\mathscr{B}_{\mathrm{NAE}})$, the latter is NP-complete. If $\mathscr{D}$ has $n \geq 2$ elements, we proceed similarly with $\mathscr{K}_n$. $\square$

Proposition 37. *In full generality, the class of problems $\{\exists, \wedge, \vee, \neq\}$-FO$(\mathscr{D})$ exhibits dichotomy: if $|D| = 1$ then the problem is in L, otherwise it is NP-complete. Consequently, the fragment extended with $=$ follows the same dichotomy.*

Proof. The proof is similar to that of Proposition 31. Let $|D| = n$. The inequality symbol $\neq$ allows to simulate $\mathscr{K}_n$. When $n = 2$, using disjunction we may simulate $\mathscr{B}_{\mathrm{NAE}}$. NP-completeness follows by reduction from CSP$(\mathscr{K}_n)$ when $n \geq 3$ and from CSP$(\mathscr{B}_{\mathrm{NAE}})$ when $n = 2$. Note that equality is not used in our hardness proof and may trivially be allowed when $|D| = 1$. Thus, the classification is the same whether one allows $=$ or not. $\square$

All fragments of the second class follow a natural dichotomy.

Corollary 38. *For any syntactic fragment $\mathscr{L}$ of FO in the second class, the model checking problem $\mathscr{L}(\mathscr{D})$ is trivial (in L) when the $\mathscr{L}$-core of $\mathscr{D}$ has one element and hard otherwise (NP-complete for existential fragments, Pspace-complete for fragments containing both quantifiers).*

3.3 Third Class

Proposition 39. *In full generality, the problem $\{\exists, \wedge, \neq\}$-FO$(\mathscr{D})$ is in L if $|D| = 1$, in P if $|D| = 2$ and $\mathscr{D}$ is bijunctive or affine, and NP-complete otherwise. The fragment extended with $=$ follows the same dichotomy.*

Proof. We classify first the fragment extended with $=$. When $|D| \geq 3$, we may use $\neq$ to simulate CSP$(\mathscr{K}_{|D|})$ which is NP-complete. When $|D| = 1$ the problem is trivially in L. We are left with the Boolean case. Let $\mathscr{D}_{\neq}$ denote the extension of $\mathscr{D}$ with $\neq$. Note that $\{\exists, \wedge, \neq\}$-FO$(\mathscr{D})$ coincides with $\{\exists, \wedge\}$-FO$(\mathscr{D}_{\neq})$ which is the Boolean CSP$(\mathscr{D}_{\neq})$. We apply Schaefer's theorem. The relation $\neq$ is neither Horn, nor dual-Horn, nor 0-valid nor 1-valid as it is not closed under any of the following Boolean operations: $\wedge$, $\vee$, c_0 or c_1 (the constant functions 0 and 1). The relation $\neq$ is both bijunctive and affine as it is closed under both the Boolean majority and minority operation (see Chen's survey for the definitions [Che09]). Consequently, $\{\exists, \wedge, \neq\}$-FO$(\mathscr{D})$ is in P if $\mathscr{D}$ is bijunctive or affine and NP-complete otherwise.

Note that we have not used $=$ in the hardness proof when $|D| \geq 3$. When $|D| = 2$, we appeal to Schaefer's theorem (Theorem 28), the proof of which relies on the Galois connection $\mathsf{Pol} - \mathsf{Inv}$ which assumes presence of $=$. However, the hardness proofs in Schaefer's theorem rely on logical reductions from $\{\exists, \wedge\}$-FO$(\mathscr{B}_{\mathrm{NAE}})$, which use definability of $\mathscr{B}_{\mathrm{NAE}}$ in $\{\exists, \wedge\}$-FO. Hence, our claim follows for the fragment $\{\exists, \wedge, \neq\}$-FO. $\square$

Proposition 40. *In full generality, the problem* $\{\exists, \forall, \wedge, \neq\}$*-FO(*$\mathscr{D}$*) is in* L *if* $|D| = 1$*, in* P *if* $|D| = 2$ *and* $\mathscr{D}$ *is bijunctive or affine, and* Pspace*-complete otherwise. The fragment extended with* = *follows the same dichotomy.*

Proof. This is similar to Proposition 39. When $|D| \geq 3$, we may use $\neq$ to simulate QCSP($\mathscr{K}_{|D|}$) which is Pspace-complete. In the Boolean case, we apply Theorem 29 to $\{\exists, \forall, \wedge\}$-FO($\mathscr{D}_{\neq}$) and the result follows.

Again equality is not used to prove hardness and the result follows for the fragment without $=$. $\qquad\square$

The case of $\{\exists, \wedge\}$-FO and $\{\exists, \wedge, =\}$-FO almost coincide as equality may be propagated out by substitution, and every sentence of the latter is logically equivalent to a sentence of the former, with the exception of sentences using only $=$ as an extensional predicate like $\exists x\, x = x$ which are tautologies as we only ever consider structures with at least one element. In the case of $\{\exists, \forall, \wedge, =\}$-FO, some equalities like $\exists x \exists y\, x = y$ and $\forall x \exists y\, x = y$ may also be propagated out by substitution. However, equalities like $\exists x \forall y\, x = y$ and $\forall x \forall y\, x = y$ can not, but they hold only in structures with a single element. This technical issue does not really affect the complexity classification, and it would suffice to consider $\{\exists, \wedge\}$-FO and $\{\exists, \forall, \wedge\}$-FO. The complexity classification for these four fragments remain open and correspond to the dichotomy conjecture for CSP and the classification program of the QCSP. In practice, we like to pretend that equality is present as it provides a better behaved algebraic framework, without affecting complexity.

This leaves the fragment $\{\exists, \forall, \wedge, \vee\}$-FO from our fourth class, which we deal with in the remainder of this paper.

4   Tetrachotomy of $\{\exists, \forall, \wedge, \vee\}$-FO($\mathscr{D}$)

The following – left as a conjecture at the end of [MM12a, MM10] – is the main contribution of this paper. Recall first that a shop f over a set D is an *A-shop* if there is an element u in D such that $f(u) = D$; and, that f is an *E-shop* if there is an element x of D such that $f^{-1}(x) = D$.

Theorem 41. *Let* $\mathscr{D}$ *be any structure.*

I. *If* $\mathscr{D}$ *is preserved by both an A-shop and an E-shop, then* $\{\exists, \forall, \wedge, \vee\}$*-FO(*$\mathscr{D}$*) is in L.*

II. *If* $\mathscr{D}$ *is preserved by an A-shop but is not preserved by any E-shop, then* $\{\exists, \forall, \wedge, \vee\}$*-FO(*$\mathscr{D}$*) is NP-complete.*

III. *If* $\mathscr{D}$ *is preserved by an E-shop but is not preserved by any A-shop, then* $\{\exists, \forall, \wedge, \vee\}$*-FO(*$\mathscr{D}$*) is co-NP-complete.*

IV. *If* $\mathscr{D}$ *is preserved neither by an A-shop nor by an E-shop, then* $\{\exists, \forall, \wedge, \vee\}$*-FO(*$\mathscr{D}$*) is Pspace-complete.*

Proof. The upper bounds (membership in L, NP and co-NP) for Cases I, II and III were known from [MM12a], but we reprove them here as a corollary of Theorem 19 to keep this paper self-contained.

Note that an A-shop is simply a U-X-shop with $U = \{u\}$, for some u in D, and $X \subseteq D$. We may therefore replace every universal quantifier by the constant u and relativise every existential quantifier to X by Theorem 19. This means that $\{\exists, \forall, \wedge, \vee\}$-FO($\mathscr{D}$) is in NP when it has an A-shop as a surjective hyper-endomorphism.

Note that an E-shop is simply a U-X-shop with $X = \{x\}$ for some x in D, and $U \subseteq D$. So Case III is dual to Case II and we finally turn to Case I.

With both an A-shop and an E-shop, we have a U-X-shop with $U = \{u\}$ and $X = \{x\}$ where u and x are in D. We may therefore replace every universal quantifier by the constant u and every existential quantifier by the constant x, by Theorem 19. We have reduced $\{\exists, \forall, \wedge, \vee\}$-FO($\mathscr{D}$) to the Boolean sentence value problem, known to be in L [Lyn77].

Theorem 46 deals with the lower bounds. NP-hardness for Case II and co-NP-hardness for Case III are proved in Subsection 4.3.3. Pspace-hardness for Case III is proved in Subsection 4.3.4. $\qquad \square$

4.1   Methodology : the Galois Connection $\mathsf{Inv} - \mathsf{shE}$

The results of this subsection ($\S$ 4.1) appeared in [MM12a] and are proved here to keep the present paper self-contained.

Let $\mathsf{shE}(\mathscr{B})$ be the set of surjective hyper-endomorphisms of $\mathscr{B}$. Let $\langle \mathscr{B} \rangle_{\{\exists, \forall, \wedge, \vee\}\text{-FO}}$ be the sets of relations that may be defined on $\mathscr{B}$ in $\{\exists, \forall, \wedge, \vee\}$-FO.

Lemma 42. *Let* $\mathbf{r} := (r_1, \ldots, r_k)$ *be a k-tuple of elements of $\mathscr{B}$. There exists a formula* $\theta_{\mathbf{r}}^{\{\exists, \forall, \wedge, \vee\}\text{-FO}}(u_1, \ldots, u_k) \in \{\exists, \forall, \wedge, \vee\}$-FO *such that the following are equivalent.*

(i) $(\mathscr{B}, r'_1, \ldots, r'_k) \models \theta_{\mathbf{r}}^{\{\exists, \forall, \wedge, \vee\}\text{-FO}}(u_1, \ldots, u_k)$.

(ii) *There is a surjective hyper-endomorphism from* $(\mathscr{B}, r_1, \ldots, r_k)$ *to* $(\mathscr{B}, r'_1, \ldots, r'_k)$.

Proof. Let $\mathbf{r} \in B^k$, $\mathbf{s} := (b_1, \ldots, b_{|B|})$ be an enumeration of B and $\mathbf{t} \in B^{|B|}$. Recall that $\varphi_{\mathscr{B}(\mathbf{r}, \mathbf{s})}(u_1, \ldots, u_k, v_1, \ldots, v_{|B|})$ is a conjunction of the positive facts of $(\mathbf{r}, \mathbf{s})$, where the variables $(\mathbf{u}, \mathbf{v})$ correspond to the elements $(\mathbf{r}, \mathbf{s})$. Similarly, $\varphi_{\mathscr{B}(\mathbf{r}, \mathbf{s}, \mathbf{t})}(u_1, \ldots, u_k, v_1, \ldots, v_{|B|}, w_1, \ldots, w_{|B|})$ is the conjunction of the positive facts of $(\mathbf{r}, \mathbf{s}, \mathbf{t})$, where the variables $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ correspond to the elements $(\mathbf{r}, \mathbf{s}, \mathbf{t})$. Set $\theta_{\mathbf{r}}^{\{\exists, \forall, \wedge, \vee\}\text{-FO}}(u_1, \ldots, u_k) :=$

$$\exists v_1, \ldots, v_{|B|} \ \varphi_{\mathscr{B}(\mathbf{r}, \mathbf{s})}(u_1, \ldots, u_k, v_1, \ldots, v_{|B|}) \wedge \forall w_1 \ldots w_{|B|}$$
$$\bigvee_{\mathbf{t} \in B^{|B|}} \varphi_{\mathscr{B}(\mathbf{r}, \mathbf{s}, \mathbf{t})}(u_1, \ldots, u_k, v_1, \ldots, v_{|B|}, w_1, \ldots, w_{|B|}).$$

[Backwards.] Suppose f is a surjective hyper-endomorphism from $(\mathscr{B}, r_1, \ldots, r_k)$ to $(\mathscr{B}', r'_1, \ldots, r'_k)$, where $\mathscr{B}' := \mathscr{B}$ (we will wish to differentiate the two occurrences of $\mathscr{B}$). We aim to prove that $\mathscr{B}' \models \theta_{\mathbf{r}}^{\{\exists, \forall, \wedge, \vee\}\text{-FO}}(r'_1, \ldots, r'_k)$. Choose arbitrary $s'_1 \in f(b_1), \ldots, s'_{|B|} \in f(b_{|B|})$ as witnesses for $v_1, \ldots, v_{|B|}$. Let $\mathbf{t}' := (t'_1, \ldots, t'_{|B|}) \in B'^{|B|}$ be any valuation of $w_1, \ldots, w_{|B|}$ and take arbitrary $t_1, \ldots, t_{|B|}$ s.t. $t'_1 \in f(t_1), \ldots, t'_{|B|} \in f(t_{|B|})$ (here we use surjectivity). Let $\mathbf{t} := (t_1, \ldots, t_{|B|})$. It follows from the definition of a surjective hyper-endomorphism that

$$\mathscr{B}' \models \varphi_{\mathscr{B}(\mathbf{r}, \mathbf{s})}(r'_1, \ldots, r'_k, s'_1, \ldots, s'_{|B|}) \wedge \varphi_{\mathscr{B}(\mathbf{r}, \mathbf{s}, \mathbf{t})}(r'_1, \ldots, r'_k, s'_1, \ldots, s'_{|B|}, t'_1, \ldots, t'_{|B|}).$$

[Forwards.] Assume that $\mathscr{B}' \models \theta_{\mathbf{r}}^{\{\exists, \forall, \wedge, \vee\}\text{-FO}}(r'_1, \ldots, r'_k)$, where $\mathscr{B}' := \mathscr{B}$. Let $b'_1, \ldots, b'_{|B|}$ be an enumeration of $B' := B$.[3] Choose some witness elements $s'_1, \ldots, s'_{|B|}$ for $v_1, \ldots, v_{|B|}$ and a witness tuple $\mathbf{t} := (t_1, \ldots, t_{|B|}) \in B^{|B|}$ s.t.

($\dagger$) $\mathscr{B}' \models \varphi_{\mathscr{B}(\mathbf{r}, \mathbf{s})}(r'_1, \ldots, r'_k, s'_1, \ldots, s'_{|B|}) \wedge \varphi_{\mathscr{B}(\mathbf{r}, \mathbf{s}, \mathbf{t})}(r'_1, \ldots, r'_k, s'_1, \ldots, s'_{|B|}, b'_1, \ldots, b'_{|B|}).$

Consider the following partial hyper-functions from B to B'.

---

[3]One may imagine $b_1, \ldots, b_{|B|}$ and $b'_1, \ldots, b'_{|B|}$ to be the same enumeration, but this is not essential. In any case, we will wish to keep the dashes on the latter set to remind us they are in $\mathscr{B}'$ and not $\mathscr{B}$.

1. $f_{\mathbf{r}}$ given by $f_{\mathbf{r}}(r_i) := \{r_i'\}$, for $1 \le i \le k$.

2. $f_{\mathbf{s}}$ given by $f_{\mathbf{s}}(b_i) = \{s_i'\}$, for $1 \le i \le |B|$.　　　　　　　　(totality)

3. $f_{\mathbf{t}}$ given by $b_i' \in f_{\mathbf{t}}(b_j)$ iff $t_i = b_j$, for $1 \le i, j \le |B|$.　　　　　　(surjectivity)

Let $f := f_{\mathbf{r}} \cup f_{\mathbf{s}} \cup f_{\mathbf{t}}$; f is a hyper-operation whose surjectivity is guaranteed by $f_{\mathbf{t}}$ (note that totality is guaranteed by $f_{\mathbf{s}}$). That f is a surjective hyper-endomorphism follows from the right-hand conjunct of (†). □

Theorem 43. *For a finite structure $\mathscr{B}$ we have* $\langle\mathscr{B}\rangle_{\{\exists,\forall,\wedge,\vee\}\text{-FO}} = \mathsf{Inv}(\mathsf{shE}(\mathscr{B}))$.

Proof.　　1. $\varphi(\mathbf{v}) \in \langle\mathscr{B}\rangle_{\{\exists,\forall,\wedge,\vee\}\text{-FO}} \Rightarrow \varphi(\mathbf{v}) \in \mathsf{Inv}(\mathsf{shE}(\mathscr{B}))$. This is proved by induction on the complexity of $\varphi(\mathbf{v})$. We only have to deal with the case of universal quantification in the inductive step, the other cases having been dealt with in the proof of the $\mathsf{Inv} - \mathsf{hE}$ Galois Connection.

  (**Inductive Step** continued from proof of Theorem 35.)

  (d) $\varphi(\mathbf{v}) := \forall w\, \psi(\mathbf{v}, w)$. Let $\mathbf{v} := (v_1, \ldots, v_l)$. Suppose $\mathscr{B} \models \forall w\, \psi(x_1, \ldots, x_l, w)$; then for each x', $\mathscr{B} \models \psi(x_1, \ldots, x_l, x')$. By IH, for any $y_1 \in f(x_1), \ldots, y_l \in f(x_l)$, we have for all y' (remember f is surjective), $\mathscr{B} \models \psi(y_1, \ldots, y_l, y')$, whereupon $\mathscr{B} \models \forall w\, \psi(y_1, \ldots, y_l, w)$.

2. $S \in \mathsf{Inv}(\mathsf{shE}(\mathscr{B})) \Rightarrow S \in \langle\mathscr{B}\rangle_{\{\exists,\forall,\wedge,\vee\}\text{-FO}}$. Consider the k-ary relation $S \in \mathsf{Inv}(\mathsf{shE}(\mathscr{B}))$. Let $\mathbf{r}_1, \ldots, \mathbf{r}_m$ be the tuples of S. Let $\theta_S^{\{\exists,\forall,\wedge,\vee\}\text{-FO}}(u_1, \ldots, u_k)$ be the following formula of $\{\exists,\forall,\wedge,\vee\}$-FO:

$$\theta_{\mathbf{r}_1}^{\{\exists,\forall,\wedge,\vee\}\text{-FO}}(u_1, \ldots, u_k) \qquad \vee \qquad \ldots \qquad \vee \qquad \theta_{\mathbf{r}_m}^{\{\exists,\forall,\wedge,\vee\}\text{-FO}}(u_1, \ldots, u_k).$$

For $\mathbf{r}_i := (r_{i1}, \ldots, r_{ik})$, note that $(\mathscr{B}, r_{i1}, \ldots, r_{ik}) \models \theta_{\mathbf{r}_i}^{\{\exists,\forall,\wedge,\vee\}\text{-FO}}(u_1, \ldots, u_k)$ (viewing the identity endomorphism as a surjective hyper endomorphism). That $\theta_S^{\{\exists,\forall,\wedge,\vee\}\text{-FO}}(u_1, \ldots, u_k) = S$ now follows from Part (ii) of Lemma 42, since $S \in \mathsf{Inv}(\mathsf{shE}(\mathscr{B}))$. □

□

Corollary 44. *Let $\mathscr{B}$ and $\mathscr{B}'$ be finite structures over the same domain B. If $\mathsf{shE}(\mathscr{B}) \subseteq \mathsf{shE}(\mathscr{B}')$ then $\{\exists,\forall,\wedge,\vee\}\text{-FO}(\mathscr{B}') \le_{\mathrm{L}} \{\exists,\forall,\wedge,\vee\}\text{-FO}(\mathscr{B})$.*

Proof. If $\mathsf{shE}(\mathscr{B}) \subseteq \mathsf{shE}(\mathscr{B}')$, then $\mathsf{Inv}(\mathsf{shE}(\mathscr{B}')) \subseteq \mathsf{Inv}(\mathsf{shE}(\mathscr{B}))$. From Theorem 43, it follows that $\langle\mathscr{B}'\rangle_{\{\exists,\forall,\wedge,\vee\}\text{-FO}} \subseteq \langle\mathscr{B}\rangle_{\{\exists,\forall,\wedge,\vee\}\text{-FO}}$. Recalling that $\mathscr{B}'$ contains only a finite number of extensional relations, we may therefore effect a Logspace reduction from $\{\exists,\forall,\wedge,\vee\}$-FO$(\mathscr{B}')$ to $\{\exists,\forall,\wedge,\vee\}$-FO$(\mathscr{B})$ by straightforward substitution of predicates. □

Consequently, the complexity of $\{\exists,\forall,\wedge,\vee\}$-FO$(\mathscr{B})$ is characterised by $\mathsf{shE}(\mathscr{B})$.

4.2　The Boolean case

We recall the case $|B| = 2$ (from [MM12a]), with the normalised domain $B := \{0,1\}$ as a warm-up. It may easily be verified that there are five DSMs in this case, depicted as a lattice in Figure 4.2. The two elements of this lattice that represent the two subgroups of S_2 are drawn in the middle and bottom. We write $\frac{0\,|\,01}{1\,|\,1}$ for the shop that sends 0 to $\{0,1\}$ and 1 to $\{1\}$.

Theorem 45 ([MM12a]). *Let $\mathscr{B}$ be a boolean structure.*

  I. *If either $\frac{0\,|\,01}{1\,|\,1}$ or $\frac{0\,|\,0}{1\,|\,01}$ is a surjective hyper-endomorphism of $\mathscr{B}$, then $\{\exists,\forall,\wedge,\vee\}$-FO$(\mathscr{B})$ is in L.*
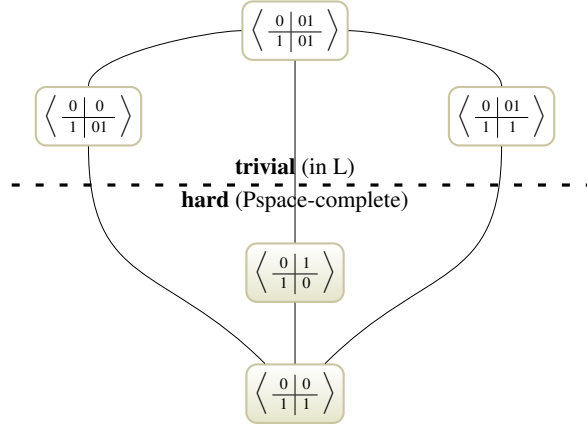
Figure 1: The boolean lattice of DSMs with their associated complexity.

II. *Otherwise, $\{\exists,\forall,\wedge,\vee\}$-FO($\mathscr{B}$) is* Pspace-*complete.*

Proof. $\mathsf{shE}(\mathscr{B})$ must be one of the five DSMs depicted in Figure 4.2. If $\mathsf{shE}(\mathscr{B})$ contains $\frac{0\,|\,01}{1\,|\,1}$ then we may relativise every existential quantifier to 1 and every universal quantifier to 0 by Theorem 18 and evaluate in L the quantifier-free part. The case of $\frac{0\,|\,0}{1\,|\,01}$ is similar with the role of 0 and 1 swapped.

We prove that if $\mathsf{shE}(\mathscr{B}) = \langle\, \frac{0\,|\,1}{1\,|\,0} \,\rangle$ then $\{\exists,\forall,\wedge,\vee\}$-FO($\mathscr{B}$) is Pspace-complete. The structure $\mathscr{K}_2$ has DSM $\mathsf{shE}(\mathscr{B}) = \langle\, \frac{0\,|\,1}{1\,|\,0} \,\rangle$. It suffices therefore to prove that $\{\exists,\forall,\wedge,\vee\}$-FO($\mathscr{K}_2$) is Pspace-hard, which we did by reduction from QCSP($\mathscr{B}_{\mathrm{NAE}}$) in the proof of Proposition 31.

If follows from Corollary 44 that when $\mathsf{shE}(\mathscr{B}) = \langle\, \frac{0\,|\,0}{1\,|\,1} \,\rangle$, $\{\exists,\forall,\wedge,\vee\}$-FO($\mathscr{B}$) is also Pspace-hard since $\langle\, \frac{0\,|\,0}{1\,|\,1} \,\rangle \subseteq \langle\, \frac{0\,|\,1}{1\,|\,0} \,\rangle$. $\qquad\square$

4.3 Proving Hardness

Our aim is to derive the following lower bounds.

Theorem 46.   II. *If $\mathscr{D}$ is preserved by an A-shop but is not preserved by any E-shop, then $\{\exists,\forall,\wedge,\vee\}$-FO($\mathscr{D}$) is NP-hard.*

   III. *If $\mathscr{D}$ is preserved by an E-shop but is not preserved by any A-shop, then $\{\exists,\forall,\wedge,\vee\}$-FO($\mathscr{D}$) is co-NP-hard.*

   IV. *If $\mathscr{D}$ is preserved neither by an A-shop nor by an E-shop, then $\{\exists,\forall,\wedge,\vee\}$-FO($\mathscr{D}$) is Pspace-hard.*

It follows from Proposition 26 and Corollary 20 that the complexity of a structure $\mathscr{D}$ is the same as the complexity of its U-X-core. Hence in this Section, *we assume w.l.o.g. that $U \cup X = D$*. We will say in this case that the DSM $\mathscr{M}$ is *reduced*. This is the critical ingredient, hitherto missing, that is needed to obtain the full classification. In order to prove Theorem 46, we need to establish the following:

   II. If U is of size one and X of size at least two then $\{\exists,\forall,\wedge,\vee\}$-FO($\mathscr{D}$) is NP-hard;

   III. If X is of size one and U of size at least two then $\{\exists,\forall,\wedge,\vee\}$-FO($\mathscr{D}$) is co-NP-hard; and,

IV. If both U and X have at least two elements then $\{\exists,\forall,\wedge,\vee\}$-FO$(\mathscr{D})$ is Pspace-hard.

In the following. we will describe a DSM $\mathscr{M}$ as being (NP-, co-NP-, Pspace-)hard in the case that $\{\exists,\forall,\wedge,\vee\}$-FO$(\mathscr{D})$ is hard for some $\mathscr{D} \in \mathsf{Inv}(\mathscr{M})$. In order to facilitate the hardness proof, we would like to show hardness of a monoid $\widehat{\mathscr{M}}$ with a very simple structure of which $\mathscr{M}$ is in fact a sub-DSM ($\widehat{\mathscr{M}}$ is the *completion* of $\mathscr{M}$). As in general $\widehat{\mathscr{M}}$ preserves fewer relations than $\mathscr{M}$, the hardness of $\mathscr{M}$ would follow. We would like the structure of $\widehat{\mathscr{M}}$ to be sufficiently simple for us to build canonically some gadgets for our hardness proof. Thus, we wish to better understand the form that elements of $\mathscr{M}$ may take. In order to do so, we first define the *canonical shop* of $\mathscr{M}$ to be the U-X shop h in $\mathscr{M}$, guaranteed by Proposition 26, with the property that $|h(z)|$ is maximal for each $z \in U \setminus X$. Note that this maximal h is unique, as given h_1 and h_2 of the form in Proposition 26, $h_1 \circ h_2$ is also of the required form, and further satisfies $|h_1 \circ h_2(z)| \geq |h_1(z)|, |h_2(z)|$, for all $z \in U \setminus X$.

4.3.1   Characterising reduced DSMs

Any U-X-shop in $\mathscr{M}$ will be shown to be in the following special form, reminiscent of the form of the canonical shop.

Definition 47. *We say that a shop f is in the* 3-permuted form *if there are a permutation ζ of $X \cap U$, a permutation χ of $X \setminus U$ and a permutation υ of $U \setminus X$ such that f satisfies:*

- *for any y in $U \cap X$, $f(y) = \{\zeta(y)\}$;*

- *for any x in $X \setminus U$, $f(x) = \{\chi(x)\}$; and,*

- *for any u in $U \setminus X$, $f(u) = \{\upsilon(u)\} \cup X_u$, where $X_u \subseteq X \setminus U$.*

Lemma 48. *If a shop f satisfies $f(X) \cap (U \setminus X) = \emptyset$ then f is in the* 3-permuted form.

Proof. The hypothesis forces an element of X to reach an element of X and Lemma 23 forces two elements of X to have different images. Since X is finite, there exists a permutation β of X such that for every x in X, $f(x) = \{\beta(x)\}$. Since Lemma 21 forces in particular an element of U to have at most one element of U in its image and since U is finite, it follows that there exists a permutation α of U such that for every u in U, $f(u) \cap U = \{\alpha(U)\}$ and $f^{-1}(u) \cap U = \{\alpha^{-1}(U)\}$.

It follows that there exists a permutation ζ of $U \cap X$ such that for any y in $U \cap X$, $f(y) = \{\zeta(y)\}$.

The existence of a permutation χ of $X \setminus U$ such that β is the disjoint union of χ and ζ follows. Hence, for any x in $X \setminus U$, $f(x) = \{\chi(x)\}$.

Similarly, there must also be a permutation υ of $U \setminus X$ such that α is the disjoint union of υ and ζ. Hence, for any u in $U \setminus X$, $f(u) \cap U = \{\upsilon(u)\}$. Elements of $U \setminus X$ may however have some images in $X \setminus U$. So we get finally that for any u in $U \setminus X$, there is some $\emptyset \subseteq X_u \subseteq X \setminus U$ such that $f(u) = \{\upsilon(u)\} \cup X_u$. This proves that f is in the 3-permuted form and we are done. $\square$

Theorem 49. *Let $\mathscr{M}$ be a reduced DSM. Every shop in $\mathscr{M}$ is in the 3-permuted form. Moreover, every U-X-shop in $\mathscr{M}$ follows the additional requirement that the elements of $U \setminus X$ cover the set $X \setminus U$, more formally that*

$$f(U \setminus X) \cap X = \bigcup_{u \in U \setminus X} X_u = X \setminus U.$$

Proof. We can now deduce easily from Lemmata 25 and 48 that U-X-shops in $\mathscr{M}$ must take the 3-permuted form. It remains to prove that an arbitrary shop f in $\mathscr{M}$ is in the 3-permuted form. Let h be the canonical shop of $\mathscr{M}$. It follows from Lemma 8 that $f' := h \circ f \circ h$ is a U-X-shop. Hence, f' is

in the 3-permuted form. Let z in X and u in $U \setminus X$. If $f(z) \ni u$ then $f'(z) \ni u$ and f' would not be in the 3-permuted form. It follows that $f(X) \cap (U \setminus X) = \emptyset$ and appealing to Lemma 48 that f is in the 3-permuted form. $\qquad\square$

We do not need the following result in order to prove our main result. But surprisingly in a reduced DSM, U and X are unique. This means that we may speak of *the canonical shop of $\mathcal{M}$* instead of some canonical U-X-shop. It also means that we can define the U-X-core of a structure $\mathcal{D}$ *without explicitly referring to U or X* as the minimal substructure of $\mathcal{D}$ which satisfy the same $\{\exists, \forall, \wedge, \vee\}$-FO sentences.

Theorem 50. *Let $\mathcal{D}$ be a structure that is both a U-X-core and a U'-X'-core then it follows that $U = U'$ and $X = X'$.*

Proof. We do a proof by contradiction. Let h and h' be the canonical U-X-shop and U'-X'-shop, respectively. Assume $U' \neq U$ and let x in $U' \setminus U$. Note that since $D = U \cup X$, our notation is consistent as x does belong to $X \setminus U$. Thus, there exists some u in $U \setminus X$ such that $h(u) \supseteq \{u, x\}$ (and necessarily $u \neq x$).

By Theorem 49, h has to be in the 3-permuted form w.r.t. U' and X', which means that h can send an element to at most one element of U'. Since x belongs to U', it follows that u belongs to $D \setminus U' = X' \setminus U'$. But the three permuted form prohibits an element of X' to reach an element of U'. A contradiction.

It does not follow yet that $X' = X$ as the pairs of sets may have shifting intersections. However, the dual argument to the above applies and yields $X = X'$. $\qquad\square$

Corollary 51. *Let $\mathcal{D}$ be a finite structure. The U-X-core of $\mathcal{D}$ is unique up to isomorphism. It is a minimal induced substructure $\widetilde{\mathcal{D}}$ of $\mathcal{D}$, that satisfies the same $\{\exists, \forall, \wedge, \vee\}$-FO formulae with free-variables in $\widetilde{D}$. Moreover, once $\widetilde{D}$ is fixed, there are two uniquely determined subsets U and X such that $U \cup X = \widetilde{\mathcal{D}} \subset D$ which are minimal within D with respect to the following equivalent properties,*

- *$\mathcal{D}$ has $\forall U$-$\exists X$-relativisation w.r.t. $\{\exists, \forall, \wedge, \vee\}$-FO; or,*

- *$\mathcal{D}$ has a U-X-shop that may act as the identity over $U \cup X$.*

Proof. The last point follows from our definition of a U-X-core and from Proposition 26. It is equivalent to the $\forall U$-$\exists X$-*relativisation* property by Theorem 19. It follows that $\mathcal{D}$ and $\widetilde{\mathcal{D}}$ satisfy the same $\{\exists, \forall, \wedge, \vee\}$-FO formulae with free-variables in $\widetilde{D}$ (see Corollary 20). Conversely, if $\mathcal{D}$ and $\widetilde{\mathcal{D}}$ satisfy the same $\{\exists, \forall, \wedge, \vee\}$-FO formulae with free-variables in $\widetilde{D}$, then $\mathcal{D}$ has $\widetilde{D}$-$\widetilde{D}$-relativisation. The existence of a "$\widetilde{D}$-$\widetilde{D}$-shop" follows by Theorem 19. Enforcing the minimality criteria, we get some U-X-shop with some $U, X \subseteq \widetilde{D}$ (this is because, we may proceed by retraction, as explained in the beginning of Subsection 2.10). Moreover, by minimality of $\widetilde{\mathcal{D}}$, we must have $U \cup X = \widetilde{D}$. We have a U-X-core as in our original definition in terms of a U-X-shop satisfying minimality criteria. It follows from Theorem 50 that U and X are unique (within $\widetilde{D}$). $\qquad\square$

Recall that the $\{\exists, \forall, \wedge, \vee\}$-FO-core $\mathcal{D}'$ of $\mathcal{D}$ is the smallest (w.r.t. domain size) structure that is $\{\exists, \forall, \wedge, \vee\}$-FO-equivalent to $\mathcal{D}$.

Proposition 52. *The notion of a U-X-core and of a $\{\exists, \forall, \wedge, \vee\}$-FO-core coincide.*

Proof. Let $\mathcal{D}$ be a structure that is a U-X-core with (unique) subsets U and X. Let c be the canonical shop of $\mathcal{D}$.

Let $\mathcal{D}'$ be a $\{\exists, \forall, \wedge, \vee\}$-FO-core of $\mathcal{D}$, that is a smallest (w.r.t. domain size) structure that is $\{\exists, \forall, \wedge, \vee\}$-FO-equivalent to $\mathcal{D}$. Let U' and X' be subsets of D' witnessing that $\mathcal{D}'$ is a U'-X' core. Note that $U' \cup X' = D'$ by minimality of $\mathcal{D}'$ (and consequently, U' and X' are uniquely determined by Theorem 50). Let c' be the canonical shop of $\mathcal{D}'$.

By Proposition 10, since $\mathscr{D}$ and $\mathscr{D}'$ are $\{\exists, \forall, \wedge, \vee\}$-FO-equivalent, there exist two surjective hypermorphisms g from $\mathscr{D}$ to $\mathscr{D}'$ and f from $\mathscr{D}'$ to $\mathscr{D}$.

Let U'' be a minimal subset of $(g)^{-1}(U')$ such that $g(U'') = U'$. Note that $f \circ c' \circ g$ is a U''-surjective shop of $\mathscr{D}$. By minimality of U, it follows that $|U| \leq |U''| \leq |U'|$. A similar argument over $\mathscr{D}'$ gives $|U'| \leq |U|$, and consequently, $|U| = |U'|$. Moreover, since $c \circ (f \circ c' \circ g)$ is a U''-surjective X-total surjective hyperendomorphism of $\mathscr{D}$, By Theorem 50, it follows that $U = U''$.

This means that there is a bijection α' from U' to U such that, for any u' in U', $g^{-1}(u') = \{\alpha'(u')\}$.

By duality we obtain similarly that $|X|=|X'|$ and that there is a bijection β from X to X' such that, for any x in X, $g(x) = \{\beta(x)\}$.

Thus, g acts necessarily as a bijection from $U \cap X$ to $U' \cap X'$.

The map $\tilde{g}$ from D to D' defined for any u in U as $\tilde{g}(u) := \alpha'^{-1}(u)$ and $\tilde{g}(x) := \beta(x)$ is a homomorphism from $\mathscr{D}$ to $\mathscr{D}'$ that is both injective and surjective.

A symmetric argument yields a map $\tilde{f}$ that is a bijective homomorphism from $\mathscr{D}'$ to $\mathscr{D}$. Isomorphism of $\mathscr{D}'$ and $\mathscr{D}$ follows.

$\square$

Remark 2. *To simplify the presentation, we defined the $\mathscr{L}$-core as a minimal structure w.r.t. domain size. Considering minimal structures w.r.t. inclusion, we would get the same notion for $\{\exists, \forall, \wedge, \vee\}$-FO. This is also the case for CSP, but it is not the case in general. For example, this is not the case for the logic $\{\exists, \forall, \wedge\}$-FO, which corresponds to QCSP [MM12b].*

Lemma 53. *Let $\mathscr{M}$ be a reduced DSM with associated sets U and X. There are only three cases possible.*

1. *$U \cap X \neq \emptyset$, $U \setminus X \neq \emptyset$ and $U \setminus X \neq \emptyset$.*

2. *$U = X$.*

3. *$U \cap X = \emptyset$.*

Proof. We prove that $U \subsetneq X$ is not possible. Otherwise, let x in $X \setminus U$ and h be the canonical shop. There exists some u in $U \subsetneq X$ such that $h(u) \ni x$ by U-surjectivity of h. Since u does not occur in the image of any other element than u under the canonical shop, this would mean that h is $X \setminus \{u\}$-total, contradicting the minimality of X.

By duality $X \subsetneq U$ is not possible either and the result follows. $\square$

4.3.2 The hard DSM above $\mathscr{M}$

Define the completion $\widehat{\mathscr{M}}$ of $\mathscr{M}$ to be the DSM that contains *all* shops in the 3-permuted form of $\mathscr{M}$. More precisely, the canonical shop of $\widehat{\mathscr{M}}$ is the shop $\hat{h}$ where every set X_u is the whole set $X \setminus U$, and, for every permutation ζ of $X \cap U$, χ of $X \setminus U$ and υ of $U \setminus X$, any shop in the 3-permuted form with these permutations is in $\widehat{\mathscr{M}}$. Note that by construction, $\mathscr{M}$ is a sub-DSM of $\widehat{\mathscr{M}}$. Note also that the minimality of U and X still holds in $\widehat{\mathscr{M}}$. We will establish hardness for $\widehat{\mathscr{M}}$, whereupon hardness of $\mathscr{M}$ follows from Theorem 43.

4.3.3 Cases II and III: NP-hardness and co-NP-hardness

We begin with Case II. We note first that $U = \{u\}$ and $|X| \geq 2$ implies $U \cap X = \emptyset$ by Lemma 53. The structure $\mathscr{K}_{|X|} \uplus \mathscr{K}_1$, the disjoint union of a clique of size $|X|$ with an isolated vertex u, has associated DSM $\widehat{\mathscr{M}}$. The problem $\{\exists, \wedge, \vee\}$-FO$(\mathscr{K}_{|X|} \uplus \mathscr{K}_1)$ is NP-hard, since the core of $\mathscr{K}_{|X|} \uplus \mathscr{K}_1$ is $\mathscr{K}_{|X|}$ by Proposition 33.

For Case III, we may assume similarly to above that $X = \{x\}$, $|U| \geq 2$ and $U \cap X = \emptyset$ by Lemma 53. We use the duality principle, which corresponds to taking the inverse of shops. Since the inverse of an $\{x\}$-total U-surjective shop with $U \geq 2$ is a $\{U\}$-total $\{x\}$-surjective shop, we may use the structure $\overline{\mathcal{K}_{|U|} \uplus \mathcal{K}_1}$ which is $\{\forall, \vee, \wedge\}$-FO-equivalent to $\overline{\mathcal{K}_{|\mathcal{U}|}}$ (and $\{\forall, \vee, \wedge\}$-FO$(\overline{\mathcal{K}_{|\mathcal{U}|}})$ is co-NP-hard).

4.3.4   case IV: Pspace-hardness

We assume that $|U| \geq 2$ and $|X| \geq 2$ and consider the tree possible cases given by Lemma 53.

Case 1: when $U \cap X \neq \emptyset$, $U \setminus X \neq \emptyset$ and $X \setminus U \neq \emptyset$ Recall that if $\mathcal{M}$ is a sub-DSM of a hard DSM $\widehat{\mathcal{M}}$ then $\mathcal{M}$ is also hard (see Theorem 43).

We write $U \Delta X$ as an abbreviation for $(X \setminus U) \cup (U \setminus X)$. To build $\widehat{\mathcal{M}}$ from $\mathcal{M}$, we added all permutations, and chose for each set $X_u = X \setminus U$. We carry on with this completion process and consider the super-DSM $\mathcal{M}'$ which is generated by a single shop g' defined as follows:

- for every y in $X \cap U$, $g'(y) := X \Delta U$; and,

- for every z in $X \Delta U$, $g'(z) := X \cap U$, where $X \Delta U$ denotes $(X \setminus U) \cup (U \setminus X)$.

The complete bipartite graph $\mathcal{K}_{X \Delta U, X \cap U}$ has $\mathcal{M}'$ for DSM. Observing that there is a full surjective homomorphism from $\mathcal{K}_{X \Delta U, X \cap U}$ to $\mathcal{K}_2$, thus by Proposition 10 the two structures agree on all sentences of $\{\exists, \forall, \wedge, \vee, \neg\}$-FO and so also on all sentences of $\{\exists, \forall, \wedge, \vee\}$-FO. It suffices therefore to prove that $\{\exists, \forall, \wedge, \vee\}$-FO$(\mathcal{K}_2)$ is Pspace-hard, which we did by reduction from QCSP$(\mathcal{B}_{\text{NAE}})$ in the proof of Theorem 45.

Case 2: when $U = X$ The clique $\mathcal{K}_{|U|}$ has DSM $\widehat{\mathcal{M}}$. The problem $\{\exists, \forall, \wedge, \vee\}$-FO$(\mathcal{K}_{|U|})$ is Pspace-complete by Theorem 45 in the Boolean case; and, beyond that, it is also Pspace-hard as a generalisation of the Pspace-complete QCSP$(\mathcal{K}_{|U|})$. The Pspace-completeness of $\{\exists, \forall, \wedge, \vee\}$-FO$(\mathcal{D})$ follows from Theorem 43.

Case 3: when $U \cap X = \emptyset$ We can no longer complete the monoid $\widehat{\mathcal{M}}$ into $\mathcal{M}'$, as we would end up with a trivial monoid. The remainder of this section is devoted to a generic hardness proof. Assume that $|U| = j \geq 2$ and $|X| = k \geq 2$ and w.l.o.g. let $U = \{1, 2, \ldots j\}$ and $X = \{j+1, j+2, \ldots j+k\}$. Recalling that the symmetric group is generated by a transposition and a cyclic permutation, let $\widehat{\mathcal{M}}$ be the DSM given by

$$
\left\langle
\begin{array}{c|c}
1 & 2,j+1,\ldots,j+k \\ \hline
2 & 1,j+1,\ldots,j+k \\ \hline
3 & 3,j+1,\ldots,j+k \\ \hline
\vdots & \vdots \\ \hline
j & j,j+1,\ldots,j+k \\ \hline
j+1 & j+1 \\ \hline
j+2 & j+2 \\ \hline
j+3 & j+3 \\ \hline
\vdots & \vdots \\ \hline
j+k & j+k
\end{array}
,
\begin{array}{c|c}
1 & 2,j+1,\ldots,j+k \\ \hline
2 & 3,j+1,\ldots,j+k \\ \hline
3 & 4,j+1,\ldots,j+k \\ \hline
\vdots & \vdots \\ \hline
j & 1,j+1,\ldots,j+k \\ \hline
j+1 & j+1 \\ \hline
j+2 & j+2 \\ \hline
j+3 & j+3 \\ \hline
\vdots & \vdots \\ \hline
j+k & j+k
\end{array}
,
\begin{array}{c|c}
1 & 1,j+1,\ldots,j+k \\ \hline
2 & 2,j+1,\ldots,j+k \\ \hline
3 & 3,j+1,\ldots,j+k \\ \hline
\vdots & \vdots \\ \hline
j & j,j+1,\ldots,j+k \\ \hline
j+1 & j+2 \\ \hline
j+2 & j+1 \\ \hline
j+3 & j+3 \\ \hline
\vdots & \vdots \\ \hline
j+k & j+k
\end{array}
,
\begin{array}{c|c}
1 & 1,j+1,\ldots,j+k \\ \hline
2 & 2,j+1,\ldots,j+k \\ \hline
3 & 3,j+1,\ldots,j+k \\ \hline
\vdots & \vdots \\ \hline
j & j,j+1,\ldots,j+k \\ \hline
j+1 & j+2 \\ \hline
j+2 & j+3 \\ \hline
j+3 & j+4 \\ \hline
\vdots & \vdots \\ \hline
j+k & j+1
\end{array}
\right\rangle .
$$

We will give a structure $\widehat{\mathcal{D}}$ such that $\mathsf{shE}(\widehat{\mathcal{D}}) = \widehat{\mathcal{M}}$. Firstly, though, given some fixed u in U and x in X, let $\mathcal{G}_{u,x}^{|U|,|X|}$ be the symmetric graph with self-loops with domain $D = U \cup X$ such that
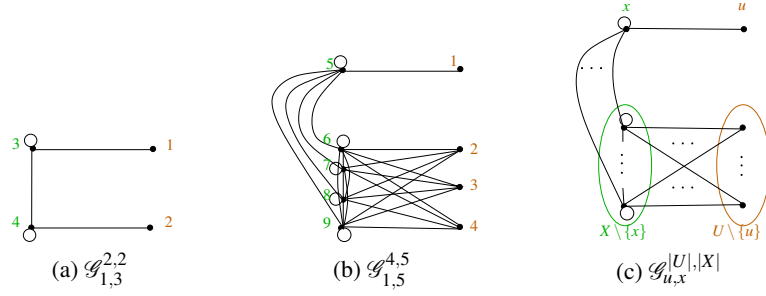
- u and x are adjacent;

(a) $\mathscr{G}_{1,3}^{2,2}$ (b) $\mathscr{G}_{1,5}^{4,5}$ (c) $\mathscr{G}_{u,x}^{|U|,|X|}$

Figure 2: Main Gadget.

- The graph induced by X is a reflexive clique $\mathscr{K}_X^{\mathrm{ref}}$; and,

- $U \setminus \{u\}$ and $X \setminus \{x\}$ are related via a complete bipartite graph $\mathscr{K}_{|X\setminus\{x\}|,|U\setminus\{u\}|}$.

The structure $\mathscr{G}_{u,x}^{|U|,|X|}$ and the more specific $\mathscr{G}_{1,5}^{4,5}$ are drawn in Figure 2. Denote by $E_{u,x}^{|U|,|X|}$ the binary relation of $\mathscr{G}_{u,x}^{|U|,|X|}$ and let $\widehat{\mathscr{D}}$ be the structure with a single 4-ary relation $R^{\widehat{\mathscr{D}}}$ with domain $\widehat{D} = U \cup X$ specified as follows,

$$R^{\widehat{\mathscr{D}}} \quad := \quad \bigcup_{u \in U}\left(\left(\bigcup_{x \in X}(u,x) \quad \times \quad E_{u,x}^{|U|,|X|}\right) \quad \cup \quad \left(\bigcup_{x_1,x_2,x_3 \in X}(x_1,x_2) \quad \times \quad E_{u,x_3}^{|U|,|X|}\right)\right).$$

Essentially, when the first argument in a quadruple is from U, then the rest of the structure allows for the unique recovery of some $\mathscr{G}_{u,x}^{|U|,|X|}$; but if the first argument is from X then all possibilities from X for the remaining arguments are allowed. In particular, we note from the last big cup that (x_1,x_2,x_3,x_4) is a tuple of $R^{\widehat{\mathscr{D}}}$ for all quadruples x_1,x_2,x_3,x_4 in X.

Lemma 54. $\mathsf{shE}(\widehat{\mathscr{D}}) = \widehat{\mathscr{M}}$.

Proof. Recall that, according to Theorem 49 and our assumption on U, X and $\widehat{\mathscr{M}}$, a maximal (w.r.t. sub-shop inclusion) shop f' is of the following form,

- for any x in $X \setminus U = X$, $f(x) = \{\chi(x)\}$; and,

- for any u in $U \setminus X = U$, $f(u) = \{\upsilon(u)\} \cup X$.

where χ and υ are permutations of X and U, respectively.

(Backwards; $\widehat{\mathscr{M}} \subseteq \mathsf{shE}(\widehat{\mathscr{D}})$.) It suffices to check that a maximal shop f' in $\widehat{\mathscr{M}}$ preserves $\widehat{\mathscr{D}}$. This holds by construction. We consider first tuples from $(x_1,x_2) \times E_{u,x_3}^{|U|,|X|}$.

- A tuple with elements from X only will map to a like tuple, which must occur, so we can ignore such tuples from now on.

- A tuple (x_1,x_2,u,x_3) maps either to $(\chi(x_1),\chi(x_2),\upsilon(u),\chi(x_3))$ which appears in $(\chi(x_1),\chi(x_2)) \times E_{\upsilon(u),\chi(x_3)}^{|U|,|X|}$, or it maps to a tuple containing only elements from X.

- A tuple (x_1,x_2,x_3,u) maps either to $(\chi(x_1),\chi(x_2),\chi(x_3),\upsilon(u))$, which appears in $(\chi(x_1),\chi(x_2)) \times E_{\upsilon(u),\chi(x_3)}^{|U|,|X|}$, or it maps to a tuple containing only elements from X.

31

We consider now tuples from $(u,x) \times E_{u,x}^{|U|,|X|}$.

- If the first coordinate u is mapped to $\upsilon(u)$, then the tuple is mapped to different tuples from $\left(\upsilon(u), \chi(x)\right) \times E_{\upsilon(u),\chi(x)}^{|U|,|X|}$, depending whether the second u is mapped to an element from X or to $\upsilon(u)$.

- Otherwise, the first coordinate u is mapped to an element x_1 from X, and some other element from u' in U occurs (or the tuple contains elements from X only) and a tuple is mapped to a tuple of the form $(x_1, \chi(x), \upsilon(u'), x_3)$ which appears in $\left(x_1, \chi(x)\right) \times E_{\upsilon(u'),x_3}^{|U|,|X|}$.

(Forwards; $\mathsf{shE}(\widehat{\mathscr{D}}) \subseteq \widehat{\mathscr{M}}$.) We proceed by contraposition, demonstrating that $R^{\widehat{\mathscr{D}}}$ is violated by any $f \notin \widehat{\mathscr{M}}$. We consider the different ways that f might not be in $\widehat{\mathscr{M}}$.

- If f is s.t. $u \in f(x)$ for $x \in X$ and $u \in U$ then we, e.g., take $(u,x,x,x) \in R^{\widehat{\mathscr{D}}}$ but $(z,u,u,u) \notin R^{\widehat{\mathscr{D}}}$ (for any $z \in f(u)$) and we are done. It follows that $f(X) = X$.

- Assume now that f is s.t. $\{x_1', x_2'\} \subseteq f(x)$ for $x_1' \neq x_2'$ and $x, x_1', x_2' \in X$. Let $u, u' \in U$ be s.t. $u' \in f(u)$. Take $(u,x,u,x) \in R^{\widehat{\mathscr{D}}}$; $(u',x_1',u',x_2') \notin R^{\widehat{\mathscr{D}}}$ and we are done. It follows that f is a permutation χ on X.

- Assume now that f is s.t. $\{u_1', u_2'\} \subseteq f(u)$ for $u_1' \neq u_2'$ and $u, u_1', u_2' \in U$. Let $x, x' \in X$ be s.t. $x' \in f(x)$. Take $(u,x,u,x) \in R^{\widehat{\mathscr{D}}}$; $(u_1',x',u_2',x') \notin R^{\widehat{\mathscr{D}}}$ and we are done. It follows that f restricted to U is a permutation υ on U.

Hence, f is a sub-shop of a maximal shop f' from the DSM $\widehat{\mathscr{M}}$, and f belongs to $\widehat{\mathscr{M}}$ (recall that a DSM is closed under sub-shops). The result follows. $\square$

Proposition 55. $\{\exists, \forall, \wedge, \vee\}$-FO$(\widehat{\mathscr{D}})$ *is Pspace-complete.*

Proof. We begin with the observation that $\{\exists, \forall, \wedge, \vee\}$-FO$(\mathscr{G}_{u,x}^{|U|,|X|})$ is Pspace-complete (for each $u \in U$ and $x \in X$). This follows straightforwardly from the Pspace-completeness of $\{\exists, \forall, \wedge, \vee\}$-FO$(\mathscr{G}_{1,3}^{2,2})$, the simplest gadget which is depicted on Figure 2a. These gadgets $\mathscr{G}_{u,x}^{|U|,|X|}$ agree on all equality-free sentences – even ones involving negation – by Proposition 10, as there is a full surjective homomorphism from $\mathscr{G}_{u,x}^{|U|,|X|}$ to $\mathscr{G}_{1,3}^{2,2}$.

We will prove that $\{\exists, \forall, \wedge, \vee\}$-FO$(\mathscr{G}_{1,3}^{2,2})$ is Pspace-hard, by reduction from the Pspace-complete problem QCSP$(\mathscr{B}_{\mathrm{NAE}})$. Recall that we may assume w.l.o.g. that universal variables are relativised to U and that existential variables are relativised to X, by Theorem 19. Let φ be an instance of QCSP$(\mathscr{B}_{\mathrm{NAE}})$. We reduce φ to a (relativised) instance ψ of $\{\exists, \forall, \wedge, \vee\}$-FO$(\mathscr{G}_{1,3}^{2,2})$. The reduction goes as follows:

- an existential variable $\exists x$ of φ is replaced by an existential variable $\exists v_x \in X$ in ψ;

- a universal variable $\forall u$ of φ is replaced by $\forall u \in U \; \exists v_u \in X, \; E(u, v_u)$ in ψ; and,

- every clause $C_i := R(\alpha, \beta, \gamma)$ in φ is replaced by the following formula in ψ,

$$\forall c_i \in U, \; E(c_i, v_\alpha) \vee E(c_i, v_\beta) \vee E(c_i, v_\gamma).$$

The truth assignment is read from $\exists$ choices in X for the variables v: we arbitrarily see one value in X as true and the other as false. It is not relevant which one is which for the problem not-all-equal satisfiability, we only need to ensure that no three variables involved in a clause can get the same value. The $\forall c_i \in U$

acts as a conjunction, enforcing "one of $v_\alpha, v_\beta, v_\gamma$ is true" and "one of $v_\alpha, v_\beta, v_\gamma$ is false". This means that at least one in three has a different value.

Now, we can prove that $\{\exists, \forall, \wedge, \vee\}$-FO($\widetilde{\mathscr{D}}$) is Pspace-complete by substituting $R(u_0, x_0, u, v)$ for each instance of $E(u, v)$ in the previous proof, and by quantifying the sentence so-produced with the prefix $\forall u_0 \in U \ \exists x_0 \in X$, once u_0 and x_0 are chosen, play proceeds as above but in the copy $\mathscr{G}_{u_0, x_0}^{|U|, |X|}$, and the result follows.   □

4.4   The Complexity of the Meta-Problem

The $\{\exists, \forall, \wedge, \vee\}$-FO($\sigma$) meta-problem takes as input a finite σ-structure $\mathscr{D}$ and answers L, NP-complete, co-NP-complete or Pspace-complete, according to the complexity of $\{\exists, \forall, \wedge, \vee\}$-FO($\mathscr{D}$). The principle result of this section is that this problem is NP-hard even for some fixed and finite signature σ_0, which consists of two binary and three unary predicates (the unaries are for convenience, but it is not clear whether a single binary suffices).

Note that one may determine if a given shop f is a surjective hyper-endomorphism of a structure $\mathscr{D}$ in, say, quadratic time in $|D|$ Since we are not interested here in distinguishing levels within P, we will henceforth consider such a test to be a basic operation. We begin with the most straightforward case.

Proposition 56. *On input $\mathscr{D}$, the question "is $\{\exists, \forall, \wedge, \vee\}$-FO-($\mathscr{D}$) in L?" is in P.*

Proof. By Theorem 41, we need to check whether there is both an A-shop and an E-shop in $\mathsf{shE}(\mathscr{D})$. In this special case, it suffices to test for each u, x in D, if the following $\{u\}$-$\{x\}$-shop f preserves $\mathscr{D}$: $f(u) := D$ and $f^{-1}(x) := D$.   □

Proposition 57. *For some fixed and finite signature σ_0, on input of a σ-structure $\mathscr{D}$, the question "is $\{\exists, \forall, \wedge, \vee\}$-FO($\mathscr{D}$) in NP (respectively, NP-complete, in co-NP, co-NP-complete)?" is NP-complete.*

Proof. The four variants are each in NP. For the first, one guesses and verifies that $\mathscr{D}$ has an A-shop, for the second, one further checks that there is no $\{u\}$-$\{x\}$-shop (see the proof of Proposition 56). Similarly for the third, one guesses and verifies that $\mathscr{D}$ has an E-shop; and, for the fourth, one further checks that there is no $\{u\}$-$\{x\}$-shop. The result then follows from Theorem 41.

For NP-hardness we will address the first problem only. The same proof will work for the second (for the third and fourth, recall that a structure $\mathscr{D}$ has an A-shop iff its complement $\overline{\mathscr{D}}$ has an E-shop). We reduce from *graph 3-colourability*. Let $\mathscr{G}$ be an undirected graph with vertices $V := \{v_1, v_2, \ldots, v_s\}$. We will build a structure $\mathscr{S}_{\mathscr{G}}$ over the domain D which consists of the disjoint union of "three colours" $\{0, 1, 2\}$, u, and the "vertices" from V.

The key observation is that there is a structure $\mathscr{G}_V$ whose class of surjective hyper-endomorphisms $\mathsf{shE}(\mathscr{G}_V)$ is generated by the following A-shop:

$$
f_V := \begin{array}{c|l}
0 & 0 \\ \hline
1 & 1 \\ \hline
2 & 2 \\ \hline
u & 0, 1, 2, u, v_1, \ldots, v_s \\ \hline
v_1 & 0, 1, 2 \\ \hline
v_2 & 0, 1, 2 \\ \hline
\vdots & \vdots \\ \hline
v_s & 0, 1, 2
\end{array}
$$

The existence of such a $\mathscr{G}_V$ is in fact guaranteed by the Galois connection, fully given in [Mar10], but that may require relations of unbounded arity, and we wish to establish our result for a fixed signature. So we will appeal to Lemma 58, below, for a σ_V-structure $\mathscr{G}_V$ with the desired class of surjective hyper-endomorphisms, where the signature σ_V consists of one binary relation and three monadic predicates. The signature σ_0 will be σ_V together with a binary relational symbol E.

The structure $\mathscr{S}_{\mathscr{G}}$ is defined as in $\mathscr{G}_V$ for symbols in σ_V, and for the additional binary symbol E, as the edge relation of the instance $\mathscr{G}$ of 3-colourability together with a clique $\mathscr{K}_3$ for the colours $\{0,1,2\}$. By construction, the following holds.

- Any surjective hyper-endomorphism g of $\mathscr{S}_{\mathscr{G}}$ will be a sub-shop of f_V.

- Restricting such a shop g to V provides a set of mutually consistent 3-colourings: i.e. we may pick arbitrarily a colour from $g(v_i)$ to get a 3-colouring $\tilde{g}$. If there is an edge between v_i and v_j in $\mathscr{G}$, then $E(v_i, v_j)$ holds in $\mathscr{S}_{\mathscr{G}}$. Since g is a shop, for any pair of colours c_i, c_j, where $c_i \in g(v_i)$ and $c_j \in g(v_j)$, we must have that $E(c_i, c_j)$ holds in $\mathscr{S}_{\mathscr{G}}$. The relation E is defined as $\mathscr{K}_3$ over the colours. Hence $c_i \neq c_j$ and we are done.

- Conversely, a 3-colouring $\tilde{g}$ induces a sub-shop g of f_V: set g as f_V over elements from $\{0,1,2,u\}$ and as $\tilde{g}$ over V. The detailed argument is similar to the above.

This proves that graph 3-colourability reduces to the meta-question "is $\{\exists, \forall, \wedge, \vee\}$-FO$(\mathscr{D})$ in NP". $\qquad\square$

Note that it follows from the given proof that the meta-problem itself is NP-hard. To see this, we take the structure $\mathscr{S}_{\mathscr{G}}$ from the proof of Proposition 57 and ask which of the four classes L, NP-complete, co-NP-complete or Pspace-complete the corresponding problem belongs to. If the answer is NP-complete then $\mathscr{G}$ was 3-colourable; otherwise the answer is Pspace-complete and $\mathscr{G}$ was not 3-colourable.

Lemma 58. *Let σ_V be a signature involving one binary relations E' and three monadic predicates* Zero, One *and* Two. *There is a σ_V-structure $\mathscr{G}_V$ such that* $\mathsf{shE}(\mathscr{G}_V) = \langle f_V \rangle$.

Proof. We begin with the graph $\mathscr{G}'$ on signature $\langle E' \rangle$, depicted on Figure 3a. Note that

$$\mathsf{shE}(\mathscr{G}) := \langle \begin{array}{c|c} c & c \\ \hline u & c,u,v \\ \hline v & c \end{array} \rangle.$$

We now replace c by $\{0,1,2\}$ and v by V to obtain a graph $\mathscr{G}''$. Formally, this graph is the unique graph
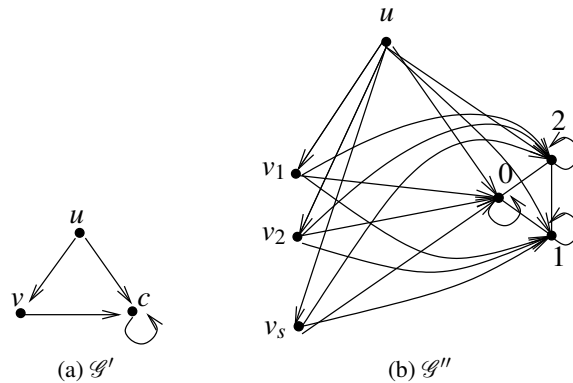

Figure 3: Building a structure with $\mathsf{shE}(\mathscr{G}_V) = \langle f_V \rangle$.

with domain $\{0,1,2,u\} \cup V$ such that the mapping which maps $\{0,1,2\}$ to c, fixes u and maps V to v, is

a strong surjective homomorphism. By construction,

$$\mathsf{shE}(\mathscr{G}'') := \big\langle \begin{array}{c|c} 0 & 0,1,2 \\ \hline 1 & 0,1,2 \\ \hline 2 & 0,1,2 \\ \hline u & 0,1,2,u,v_1,\ldots,v_s \\ \hline v_1 & 0,1,2 \\ \hline \vdots & \vdots \\ \hline v_s & 0,1,2 \end{array} \big\rangle.$$

We now build $\mathscr{G}_V$ as the structure with binary relation E' which is the edge relation from $\mathscr{G}''$ and by setting the unary predicates as follows: Zero holds only over 0, One holds only over 1 and Two holds only over 2. This effectively fixes surjective hyper-endomorphisms to act as the identity over the colours $\{0,1,2\}$ as required. □

5 Conclusion

We have classified the complexity of the model checking problem for all fragments of FO but those corresponding to the CSP and the QCSP. Our results are summarised as Figure 4.4. The inclusion of fragments is denoted by dashed edges, a larger fragment being above. Each fragment is classified in two fashions. Firstly, we have indicated on the figure the notion of core used to classify fragments, by regrouping them in the same box. Secondly, we have organised the fragments in four classes according to the nature of the complexity classification they follow. The first class is trivial. Tractability for a fragment $\mathscr{L}$ of the second class corresponds precisely to having a one element $\mathscr{L}$-core. The third class regroups fragments which have a non trivial classification viz complexity, in the sense that it does not always depend on the size of the $\mathscr{L}$-core, and include the two open cases of CSP and QCSP which we discuss in some detail below. The fourth class contains the fragment $\{\exists,\forall,\wedge,\vee\}$-FO which exhibits a behaviour intermediate between the third class and the fourth class: its complexity is fully explained in terms of the U-X-core, yet as this notion involves two sets, the fragment exhibits richness in its ensuing tetrachotomy.

For the CSP, the dichotomy conjecture has been proved in the Boolean case by Schaefer (see Theorem 28) and in the case of undirected graphs.

Theorem 59 ([HN90]). *Let $\mathscr{G}$ be an undirected graph. If $\mathscr{G}$ is bipartite then $CSP(\mathscr{G})$ is in* L, *otherwise $CSP(\mathscr{G})$ is* NP-*complete.*[4]

For CSP in general, it would suffice to settle the dichotomy conjecture for (certain) directed graphs [FV98]. The dichotomy conjecture has been settled for smooth digraphs (graphs with no sources and no sinks) [BKN09]. According to the algebraic reformulation of the dichotomy conjecture, it would suffice to prove that every structure that has a *Sigger's term* has a tractable CSP (see [BV08, Bul11] for recent surveys on the algebraic approach to the dichotomy conjecture).

For the QCSP, much less is known. We have already seen that a dichotomy between P and Pspace-complete holds in the Boolean case (Theorem 29). However, the complexity is not even known for undirected graphs. It is fully classified for graphs with at most one cycle.

Theorem 60 ([MM06]). *Let $\mathscr{G}$ be an undirected graph.*

- *If $\mathscr{G}$ is bipartite then $QCSP(\mathscr{G})$ is in* L*;*

---

[4]In the bipartite case, assuming that the graph $\mathscr{G}$ has at least one edge, then the core of $\mathscr{G}$ is $\mathscr{K}_2$. The problem $CSP(\mathscr{K}_2)$ is 2-colourability which is in the complexity class *symmetric logspace* now known to be equal to L [Rei08].
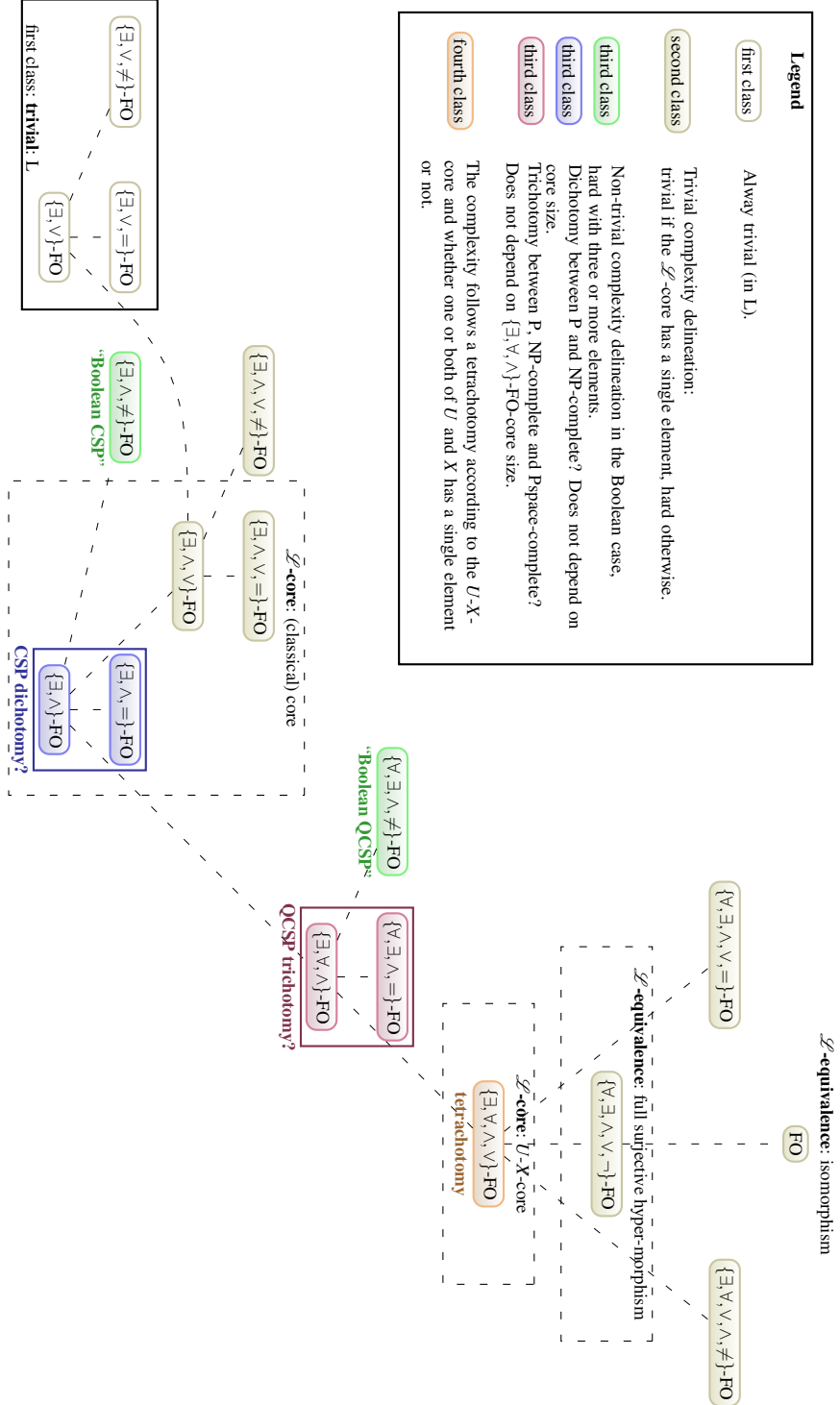
Figure 4: Classification of the complexity of the model-checking problem

- *if $\mathcal{G}$ is not bipartite and not connected then QCSP($\mathcal{G}$) is* NP*-complete; and,*

- *if $\mathcal{G}$ not bipartite, connected and contains at most one cycle then QCSP($\mathcal{G}$) is* Pspace*-complete.*

The algebraic approach to QCSP uses *surjective polymorphisms* and has led to a trichotomy in the case where all graphs of permutations are available. Recall first the definition of some special surjective operations. A k-ary *near-unanimity operation f* satisfies

$$f(x_1,\ldots,x_k) = \begin{cases} x & \text{if } \{x_1,\ldots,x_k\} = \{x\}; \text{ and,} \\ x & \text{if all but one of } x_1,\ldots,x_k \text{ is equal to } x. \end{cases}$$

When $k = 3$, we speak of a *majority operation*. The *k-ary near projection operation* is defined as

$$l_k(x_1,\ldots,x_k) = \begin{cases} x_1 & \text{when } |\{x_1,\ldots,x_k\}| = k; \text{ and,} \\ x_k & \text{otherwise.} \end{cases}$$

The *ternary switching operation* is defined as

$$s(x,y,z) = \begin{cases} x & \text{if } y = z, \\ y & \text{if } x = z, \\ z & \text{otherwise.} \end{cases}$$

The *dual discriminator operation* is defined as

$$d(x,y,z) = \begin{cases} y & \text{if } y = z; \text{ and,} \\ x & \text{otherwise.} \end{cases}$$

When $f(x,y,z) = x - y + z$ w.r.t. some Abelian group structure, we say that f is an *affine operation*.

Theorem 61 ([BBC$^+$09]). *Let $\mathcal{D}$ be a structure such that there is an extensional binary symbol for each graph of a permutation of D. Then the complexity of QCSP($\mathcal{D}$) follows the following trichotomy.*

- *If $\mathcal{D}$ has a surjective polymorphism which is the dual discriminator, the switching operation or an affine operation then QCSP($\mathcal{D}$) is* P.

- *Else, if $l_{|D|}$ is a surjective polymorphism of $\mathcal{D}$ then QCSP($\mathcal{D}$) is* NP*-complete.*

- *Otherwise, QCSP($\mathcal{D}$) is* Pspace*-complete.*

In general, it is known that if a structure $\mathcal{D}$ is preserved by a *near-unanimity operation* then QCSP($\mathcal{D}$) is in P, because it implies a property of *collapsibility*. This property means that an instance holds if, and only, if all sentences induced by keeping only a bounded number of universal quantifiers – the so-called *collapsings* – hold [Che08].

For undirected partially reflexive graphs (*i.e.* with possible self-loops), we have the following partial classification (reformulated algebraically).

Theorem 62 ([Mar11]). *Let $\mathcal{T}$ be a partially reflexive forest.*

- *If $\mathcal{T}$ is $\{\exists,\forall,\wedge\}$-FO-equivalent to a structure that is preserved by a majority operation then QCSP($\mathcal{T}$) is* P*; and,*

- *otherwise, QCSP($\mathcal{T}$) is* NP*-hard.*

In the case of structures with all constants, Hubie Chen has ventured some conjecture regarding the NP/Pspace-hard border: he suggests that the *polynomially generated power property (PGP)* – a property which generalises collapsibility – explains a drop in complexity to NP(see [Che12] for details).

Acknowledgment

References

[BBC$^+$09] F. Börner, A. A. Bulatov, H. Chen, P. Jeavons, and A. A. Krokhin. The complexity of constraint satisfaction games and QCSP. *Inf. Comput.*, 207(9):923–944, 2009.

[BHR09] Manuel Bodirsky, Miki Hermann, and Florian Richoux. Complexity of existential positive first-order logic. In Klaus Ambos-Spies, Benedikt Löwe, and Wolfgang Merkle, editors, *CiE*, volume 5635 of *Lecture Notes in Computer Science*, pages 31–36. Springer, 2009.

[BJK05] Andrei A. Bulatov, Peter Jeavons, and Andrei A. Krokhin. Classifying the complexity of constraints using finite algebras. *SIAM J. Comput.*, 34(3):720–742, 2005.

[BKJ00] Andrei A. Bulatov, Andrei A. Krokhin, and Peter Jeavons. Constraint satisfaction problems and finite algebras. In Ugo Montanari, José D. P. Rolim, and Emo Welzl, editors, *ICALP*, volume 1853 of *Lecture Notes in Computer Science*, pages 272–282. Springer, 2000.

[BKN09] Libor Barto, Marcin Kozik, and Todd Niven. The CSP dichotomy holds for digraphs with no sources and no sinks (a positive answer to a conjecture of bang-jensen and hell). *SIAM J. Comput.*, 38(5):1782–1802, 2009.

[Bör00] Ferdinand Börner. Total multifunctions and relations. In *AAA60: Workshop on General Algebra, Dresden, Germany*, 2000.

[Bör08] Ferdinand Börner. Basics of galois connections. In Creignou et al. [CKV08], pages 38–67.

[Bul06] Andrei A. Bulatov. A dichotomy theorem for constraint satisfaction problems on a 3-element set. *J. ACM*, 53(1):66–120, 2006.

[Bul11] Andrei A. Bulatov. On the CSP dichotomy conjecture. In Alexander S. Kulikov and Nikolay K. Vereshchagin, editors, *CSR*, volume 6651 of *Lecture Notes in Computer Science*, pages 331–344. Springer, 2011.

[BV08] A. A. Bulatov and M. Valeriote. Recent results on the algebraic approach to the CSP. In Creignou et al. [CKV08], pages 68–92.

[Che04] Hubie Chen. Quantified constraint satisfaction and 2-semilattice polymorphisms. In Mark Wallace, editor, *CP*, volume 3258 of *Lecture Notes in Computer Science*, pages 168–181. Springer, 2004.

[Che08] Hubie Chen. The complexity of quantified constraint satisfaction: Collapsibility, sink algebras, and the three-element case. *SIAM J. Comput.*, 37(5):1674–1701, 2008.

[Che09] H. Chen. A rendezvous of logic, complexity, and algebra. *ACM Comput. Surv.*, 42(1), 2009.

[Che12] Hubie Chen. Meditations on quantified constraint satisfaction. *CoRR*, abs/1201.6306, 2012.

[CKS01] N. Creignou, S. Khanna, and M. Sudan. *Complexity classifications of boolean constraint satisfaction problems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001.

[CKV08] Nadia Creignou, Phokion G. Kolaitis, and Heribert Vollmer, editors. *Complexity of Constraints - An Overview of Current Research Themes [Result of a Dagstuhl Seminar]*, volume 5250 of *Lecture Notes in Computer Science*. Springer, 2008.

[CM77] Ashok K. Chandra and Philip M. Merlin. Optimal implementation of conjunctive queries in relational data bases. In John E. Hopcroft, Emily P. Friedman, and Michael A. Harrison, editors, *STOC*, pages 77–90. ACM, 1977.

[Dal97] Victor Dalmau. Some dichotomy theorems on constant-free quantified boolean formulas. Technical Report LSI-97-43-R., Departament LSI, Universitat Pompeu Fabra., 1997.

[FV98] Tomás Feder and Moshe Y. Vardi. The computational structure of monotone monadic snp and constraint satisfaction: A study through datalog and group theory. *SIAM J. Comput.*, 28(1):57–104, 1998.

[GKL+07]   Erich Grädel, P. G. Kolaitis, L. Libkin, M. Marx, J. Spencer, Moshe Y. Vardi, Y. Venema, and Scott Weinstein. *Finite Model Theory and Its Applications (Texts in Theoretical Computer Science. An EATCS Series)*. Springer, 2007.

[HN90]   P. Hell and J. Nešetřil. On the complexity of H-coloring. *J. Combin. Theory Ser. B*, 48, 1990.

[HR09]   Miki Hermann and Florian Richoux. On the computational complexity of monotone constraint satisfaction problems. In Sandip Das and Ryuhei Uehara, editors, *WALCOM*, volume 5431 of *Lecture Notes in Computer Science*, pages 286–297. Springer, 2009.

[JCG97]   Peter Jeavons, David A. Cohen, and Marc Gyssens. Closure properties of constraints. *J. ACM*, 44(4):527–548, 1997.

[Kra38]   M. Krasner. Une généralisation de la notion de corps. *Journal de Mathématiques Pures et Appliquées*, 9:367–385, 1938.

[KV00]   Phokion G. Kolaitis and Moshe Y. Vardi. Conjunctive-query containment and constraint satisfaction. *J. Comput. Syst. Sci.*, 61(2):302–332, 2000.

[Lyn77]   Nancy A. Lynch. Log space recognition and translation of parenthesis languages. *J. ACM*, 24(4):583–590, 1977.

[Mar06]   Barnaby Martin. Dichotomies and duality in first-order model checking problems. *CoRR*, abs/cs/0609022, 2006.

[Mar08]   Barnaby Martin. First-order model checking problems parameterized by the model. In Arnold Beckmann, Costas Dimitracopoulos, and Benedikt Löwe, editors, *CiE*, volume 5028 of *Lecture Notes in Computer Science*, pages 417–427. Springer, 2008.

[Mar10]   Barnaby Martin. The lattice structure of sets of surjective hyper-operations. In David Cohen, editor, *CP*, volume 6308 of *Lecture Notes in Computer Science*, pages 368–382. Springer, 2010.

[Mar11]   Barnaby Martin. QCSP on partially reflexive forests. In Jimmy Ho-Man Lee, editor, *CP*, volume 6876 of *Lecture Notes in Computer Science*, pages 546–560. Springer, 2011.

[MM06]   Barnaby Martin and Florent R. Madelaine. Towards a trichotomy for quantified *h*-coloring. In Arnold Beckmann, Ulrich Berger, Benedikt Löwe, and John V. Tucker, editors, *CiE*, volume 3988 of *Lecture Notes in Computer Science*, pages 342–352. Springer, 2006.

[MM09]   Florent R. Madelaine and Barnaby Martin. The complexity of positive first-order logic without equality. In *LICS*, pages 429–438. IEEE Computer Society, 2009.

[MM10]   Barnaby Martin and Jos Martin. The complexity of positive first-order logic without equality II: The four-element case. In Anuj Dawar and Helmut Veith, editors, *CSL*, volume 6247 of *Lecture Notes in Computer Science*, pages 426–438. Springer, 2010.

[MM11]   Florent R. Madelaine and Barnaby Martin. A tetrachotomy for positive first-order logic without equality. In Martin Grohe, editor, *LICS*, pages 311–320. IEEE Computer Society, 2011.

[MM12a]   Florent R. Madelaine and Barnaby Martin. The complexity of positive first-order logic without equality. *ACM Trans. Comput. Log.*, 13(1):5, 2012.

[MM12b]   Florent R. Madelaine and Barnaby Martin. Containment, equivalence and coreness from CSP to QCSP and beyond. *CoRR*, abs/1204.5981, 2012. To appear at CP 2012.

[Rei08]   Omer Reingold. Undirected connectivity in log-space. *J. ACM*, 55(4), 2008.

[Sch78]   T.J. Schaefer. The complexity of satisfiability problems. In *STOC*, 1978.

[Var82]   Moshe Y. Vardi. The complexity of relational query languages (extended abstract). In Harry R. Lewis, Barbara B. Simons, Walter A. Burkhard, and Lawrence H. Landweber, editors, *STOC*, pages 137–146. ACM, 1982.